OM protein - protein search, using sw model

Run on: October 5, 2004, 06:07:39 ; Search time 9.16981 Seconds
(without alignments)
94.410 Million cell updates/sec

Title: US-10-022-286-1
Perfect score: 37
Sequence: 1 LAASLLSRV 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	37	100.0	637	2	D70888	hypothetical prote
2	31	83.8	1777	2	T00490	nonstructural prot
3	30	81.1	290	2	E70703	probable uspA prot
4	30	81.1	422	2	H64600	UDP-N-acetylglucos
5	30	81.1	1086	2	S74251	phosphorylase kina
6	30	81.1	9376	2	T14593	syringomycin synth
7	29	78.4	186	2	T32528	hypothetical prote
8	29	78.4	438	2	T35355	hypothetical prote
9	29	78.4	832	2	A40205	Na+/H+-exchanging
10	29	78.4	865	2	A53186	fluG protein - Eme
11	29	78.4	889	2	T11742	egg sperm receptor
12	29	78.4	1278	1	E69615	enterobactin synth
13	28	75.7	114	2	E95195	hypothetical prote
14	28	75.7	146	2	F72709	hypothetical prote
15	28	75.7	198	2	C97638	hypothetical prote
16	28	75.7	198	2	AD2861	hypothetical prote
17	28	75.7	257	2	C81999	parA family protei
18	28	75.7	257	2	D81227	ParA family protei
19	28	75.7	263	2	T38351	ribosomal protein
20	28	75.7	303	2	A84752	hypothetical prote
21	28	75.7	312	2	B90248	deoxyhypusine synt
22	28	75.7	421	2	H72492	probable proton/so
23	28	75.7	431	2	F95236	PTS system, IIC co
24	28	75.7	431	2	H98100	hypothetical prote
25	28	75.7	450	2	A38171	L-lysine 6-transam
26	28	75.7	493	2	A65126	probable general s
27	28	75.7	497	2	C34469	pulE protein - Kle
28	28	75.7	519	2	F86265	hypothetical prote
29	28	75.7	531	2	T43551	multidrug resistan
30	28	75.7	534	2	T30629	hypothetical prote
31	28	75.7	567	2	A83292	probable ATP-depen
32	28	75.7	572	2	H96685	probable AMP-bindi
33	28	75.7	583	2	D82672	general secretory
34	28	75.7	676	2	AB2017	two-component sens
35	28	75.7	802	2	T23584	hypothetical prote
36	28	75.7	912	2	T18785	hypothetical prote
37	28	75.7	1046	2	H87318	TonB-dependent rec
38	28	75.7	1468	1	S30818	hypothetical prote
39	28	75.7	1601	2	T18800	hypothetical prote
40	28	75.7	2327	2	T31733	hypothetical prote
41	28	75.7	2649	2	A40937	bullous pemphigoid
42	28	75.7	3144	2	S64791	VPS13 protein - ye
43	27	73.0	120	2	B82963	hypothetical prote
44	27	73.0	136	2	JS0740	H+-transporting tw
45	27	73.0	194	2	A83371	probable transcrip

ALIGNMENTS

RESULT 1
D70888
hypothetical protein Rv3808c - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C;Accession: D70888
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, ; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: D70888
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-637 <COL>
A;Cross-references: GB:AL022076; GB:AL123456; NID:g3256026; PIDN:CAA17872.1; PID:g29504
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: Rv3808c
C;Superfamily: Mycobacterium tuberculosis hypothetical protein Rv3808c

Query Match 100.0%; Score 37; DB 2; Length 637;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LAASLLSRV 9
|||||||||
Db 4 LAASLLSRV 12

RESULT 2
T00490
nonstructural protein precursor - himetobi P virus
C;Species: himetobi P virus
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 08-Oct-1999
C;Accession: T00490
R;Nakashima, N.; Sasaki, J.
submitted to the EMBL Data Library, August 1998
A;Description: Complete nucleotide sequence of an insect picorna-like virus, Himetobi P
A;Reference number: Z14156
A;Accession: T00490
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: genomic RNA
A;Residues: 1-1777 <NAK>
A;Cross-references: EMBL:AB017037; NID:d1226972; PIDN:BAA32553.1; PID:d1033516
C;Keywords: nonstructural protein

Query Match 83.8%; Score 31; DB 2; Length 1777;
Best Local Similarity 66.7%; Pred. No. 1.8e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 LAASLLSRV 9
              ||||:||::
Db        587 LAASILSKI 595
```

RESULT 3
E70703
probable uspA protein - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C;Accession: E70703
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: E70703
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-290 <COL>
A;Cross-references: GB:Z79702; GB:AL123456; NID:g3261642; PIDN:CAB02051.1; PID:g1524284
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: uspA
C;Superfamily: inner membrane protein ugpA

  Query Match   81.1%; Score 30; DB 2; Length 290;
  Best Local Similarity   87.5%; Pred. No. 47;
  Matches   7; Conservative   1; Mismatches   0; Indels   0; Gaps   0;

```
Qy          1 LAASLLSR 8
              ||||||:|
Db         89 LAASLLAR 96
```

RESULT 4
H64600
UDP-N-acetylglucosamine enolpyruvyl transferase - Helicobacter pylori (strain 26695)
C;Species: Helicobacter pylori
C;Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 29-Sep-1999
C;Accession: H64600
R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.
 Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenne
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.
Nature 388, 539-547, 1997
A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A;Reference number: A64520; MUID:97394467; PMID:9252185
A;Accession: H64600
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-422 <TOM>
A;Cross-references: GB:AE000578; GB:AE000511; NID:g2313759; PIDN:AAD07708.1; PID:g231376
C;Genetics:
A;Start codon: TTG
C;Superfamily: UDP-N-acetylglucosamine 1-carboxyvinyltransferase MurZ

  Query Match   81.1%; Score 30; DB 2; Length 422;
  Best Local Similarity   87.5%; Pred. No. 69;
  Matches   7; Conservative   1; Mismatches   0; Indels   0; Gaps   0;

```
Qy          1 LAASLLSR 8
              |||:||||
Db         29 LAATLLSR 36
```

RESULT 5
S74251
phosphorylase kinase (EC 2.7.1.38) beta chain - human
C;Species: Homo sapiens (man)
C;Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 09-Dec-2002
C;Accession: S74251
R;Wuellrich-Schmoll, A.; Kilimann, M.W.
Eur. J. Biochem. 238, 374-380, 1996
A;Title: Structure of the human gene encoding the phosphorylase kinase beta subunit (PHK
A;Reference number: S74249; MUID:96283831; PMID:8681948
A;Accession: S74251
A;Molecule type: DNA
A;Residues: 1-1086 <WUE>
A;Cross-references: EMBL:X84937
A;Experimental source: brain
C;Genetics:
A;Gene: GDB:PHKB
A;Cross-references: GDB:120286; OMIM:172490
A;Map position: 16q12.1-16q12.1
A;Introns: 19/1; 49/1; 95/2; 128/3; 164/3; 191/3; 230/2; 251/3; 283/3; 349/3; 369/1; 39
8/3; 994/3; 1041/3
C;Superfamily: phosphorylase kinase beta chain
C;Keywords: alternative splicing; autophosphorylation; lipoprotein; muscle; phosphoprot
F;12,20,694/Binding site: phosphate (Ser) (covalent) (by autophosphorylation) #status p
F;1083/Binding site: farnesyl (Cys) (covalent) #status predicted

  Query Match   81.1%; Score 30; DB 2; Length 1086;
  Best Local Similarity   87.5%; Pred. No. 1.8e+02;
  Matches   7; Conservative   1; Mismatches   0; Indels   0; Gaps   0;

```
Qy          2 AASLLSRV 9
              ||||||:|
Db        778 AASLLSKV 785
```

RESULT 6
T14593
syringomycin synthetase - Pseudomonas syringae pv. syringae
C;Species: Pseudomonas syringae pv. syringae
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 01-Dec-2000
C;Accession: T14593
R;Guenzi, E.; Galli, G.; Grgurina, I.; Gross, D.C.; Grandi, G.
J. Biol. Chem. 273, 32857-32863, 1998
A;Title: Characterization of the syringomycin synthetase gene cluster. A link between p
A;Reference number: Z18153; MUID:99047670; PMID:9830033
A;Accession: T14593
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-9376 <GUE>
A;Cross-references: EMBL:AF047828; NID:g3510628; PID:g3510629; PIDN:AAC80285.1
C;Genetics:
A;Gene: syrE
C;Superfamily: acetate-CoA ligase homology; acyl carrier protein homology
C;Keywords: carrier protein; phosphopantetheine; phosphoprotein
F;441-896/Domain: acetate-CoA ligase homology <ACL1>
F;914-981/Domain: acyl carrier protein homology <ACP1>
F;1529-1984/Domain: acetate-CoA ligase homology <ACL2>
F;2002-2069/Domain: acyl carrier protein homology <ACP2>
F;2613-3071/Domain: acetate-CoA ligase homology <ACL3>
F;3089-3156/Domain: acyl carrier protein homology <ACP3>
F;3700-4158/Domain: acetate-CoA ligase homology <ACL4>
F;4176-4244/Domain: acyl carrier protein homology <ACP4>
F;4768-5224/Domain: acetate-CoA ligase homology <ACL5>
F;5242-5310/Domain: acyl carrier protein homology <ACP5>
F;5834-6280/Domain: acetate-CoA ligase homology <ACL6>
F;6298-6366/Domain: acyl carrier protein homology <ACP6>
F;6892-7352/Domain: acetate-CoA ligase homology <ACL7>
F;7370-7437/Domain: acyl carrier protein homology <ACP7>
F;7990-8440/Domain: acetate-CoA ligase homology <ACL8>
F;8458-8526/Domain: acyl carrier protein homology <ACP8>
F;9015-9083/Domain: acyl carrier protein homology <ACP9>
F;946,2034,3121,4208,5274,6330/Binding site: phosphopantetheine (Ser) (covalent) #stat

  Query Match   81.1%; Score 30; DB 2; Length 9376;
  Best Local Similarity   77.8%; Pred. No. 1.7e+03;
  Matches   7; Conservative   1; Mismatches   1; Indels   0; Gaps   0;

```
Qy          1 LAASLLSRV 9
              ||| |:|||
Db       6332 LAAQLVSRV 6340
```

RESULT 7
T32528
hypothetical protein CD4.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 04-Mar-2000
C;Accession: T32528
R;Du, Z.; Scheet, P.
submitted to the EMBL Data Library, December 1997
A;Description: The sequence of C. elegans cosmid CD4.
A;Reference number: Z21185
A;Accession: T32528
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-186 <DUZ>
A;Cross-references: EMBL:AF036694; PIDN:AAB88341.1; GSPDB:GN00022; CESP:CD4.3
A;Experimental source: strain Bristol N2; clone CD4
C;Genetics:
A;Gene: CESP:CD4.3
A;Map position: 4
A;Introns: 47/2; 143/3
C;Superfamily: Caenorhabditis elegans hypothetical protein CD4.3

```
  Query Match             78.4%;  Score 29;  DB 2;  Length 186;
  Best Local Similarity   87.5%;  Pred. No. 50;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 AASLLSRV 9
              | ||||||
Db          5 AVSLLSRV 12
```

RESULT 8
T35355
hypothetical protein SC5H1.39 - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
C;Accession: T35355
R;Oliver, K.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, May 1999
A;Reference number: Z21575
A;Accession: T35355
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-438 <OLI>
A;Cross-references: EMBL:AL049863; PIDN:CAB42964.1; GSPDB:GN00070; SCOEDB:SC5H1.39
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SC5H1.39

```
  Query Match             78.4%;  Score 29;  DB 2;  Length 438;
  Best Local Similarity   87.5%;  Pred. No. 1.2e+02;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 AASLLSRV 9
              | ||||||
Db         84 AGSLLSRV 91
```

RESULT 9
A40205
Na+/H+-exchanging protein 3 - rabbit
N;Alternate names: Na+/H+ antiporter
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 05-Nov-1999
C;Accession: A40205
R;Tse, C.M.; Brant, S.R.; Walker, M.S.; Pouyssegur, J.; Donowitz, M.
J. Biol. Chem. 267, 9340-9346, 1992
A;Title: Cloning and sequencing of a rabbit cDNA encoding an intestinal and kidney-speci
A;Reference number: A40205; MUID:92250540; PMID:1374392
A;Accession: A40205
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-832 <TSE>
A;Cross-references: GB:M87007; NID:g165548; PIDN:AAA31420.1; PID:g165549
C;Keywords: transmembrane protein

```
  Query Match             78.4%;  Score 29;  DB 2;  Length 832;
  Best Local Similarity   77.8%;  Pred. No. 2.4e+02;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 LAASLLSRV 9
              || |||:||
Db        766 LAPSLLARV 774
```

RESULT 10
A53186
fluG protein - Emericella nidulans
C;Species: Emericella nidulans, Aspergillus nidulans
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 07-May-1999
C;Accession: A53186
R;Lee, B.N.; Adams, T.H.
Genes Dev. 8, 641-651, 1994
A;Title: The Aspergillus nidulans fluG gene is required for production of an extracellu
A;Reference number: A53186; MUID:95011568; PMID:7926755
A;Accession: A53186
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-865 <LEE>
A;Cross-references: GB:L27817; NID:g450257; PID:g450258
C;Genetics:
A;Gene: fluG
A;Introns: 375/3; 790/1

```
  Query Match             78.4%;  Score 29;  DB 2;  Length 865;
  Best Local Similarity   66.7%;  Pred. No. 2.5e+02;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LAASLLSRV 9
              ||||:||::
Db        140 LAASVLSQI 148
```

RESULT 11
T11742
egg sperm receptor - sea urchin (Strongylocentrotus purpuratus)
C;Species: Strongylocentrotus purpuratus (purple urchin)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 24-Nov-1999
C;Accession: T11742
R;Foltz, K.R.; Partin, J.S.; Lennarz, W.J.
Science 259, 1421-1425, 1993
A;Title: Sea urchin egg receptor for sperm: sequence similarity of binding domain and hs
A;Reference number: Z17324; MUID:93197888; PMID:8383878
A;Accession: T11742
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-889 <FOL>
A;Cross-references: EMBL:L04969; NID:g1580782; PID:g1580783
A;Experimental source: immature ovary
C;Superfamily: heat shock protein 91

```
  Query Match             78.4%;  Score 29;  DB 2;  Length 889;
  Best Local Similarity   77.8%;  Pred. No. 2.5e+02;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 LAASLLSRV 9
              ||| || ||
Db        309 LAAELLKRV 317
```

RESULT 12

E69615
enterobactin synthetase component dhbF - Bacillus subtilis
N;Alternate names: enterochelin synthetase component F
N;Contains: (2,3-dihydroxybenzoyl)serine O-[(2,3-dihydroxybenzoyl)seryl]transferase (EC
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 10-Jul-1998 #text_change 03-Nov-2000
C;Accession: E69615
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: E69615
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1278 <KUN>
A;Cross-references: GB:Z99120; GB:AL009126; NID:g2635613; PIDN:CAB15186.1; PID:g2635693
A;Experimental source: strain 168
C;Genetics:
A;Gene: dhbF
C;Function:
A;Description: forms enteroabactin from three molecules of (2,3-dihydroxybenzoyl)serine
er protein
A;Pathway: enterobactin biosynthesis
C;Superfamily: enterobactin synthetase component F; acetate-CoA ligase homology; acyl ca
C;Keywords: acyltransferase; carrier protein; enterobactin biosynthesis; iron transport;
F;504-948/Domain: acetate-CoA ligase homology <ACL>
F;964-1032/Domain: acyl carrier protein homology <ACP>
F;996/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 78.4%; Score 29; DB 1; Length 1278;
Best Local Similarity 66.7%; Pred. No. 3.7e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LAASLLSRV 9
||| |:||:
Db 998 LAARLMSRI 1006

RESULT 13
E95195
hypothetical protein SP1679 [imported] - Streptococcus pneumoniae (strain TIGR4)
C;Species: Streptococcus pneumoniae
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C;Accession: E95195
R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A;Reference number: A95000; MUID:21357209; PMID:11463916
A;Accession: E95195
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-114 <KUR>
A;Cross-references: GB:AE005672; PIDN:AAK75758.1; PID:g14973172; GSPDB:GN00164; TIGR:SP4
A;Experimental source: strain TIGR4
C;Genetics:
A;Gene: SP1679

Query Match 75.7%; Score 28; DB 2; Length 114;
Best Local Similarity 66.7%; Pred. No. 51;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LAASLLSRV 9
:| |:|||
Db 52 MARSILSRV 60

RESULT 14
F72709
hypothetical protein APE1093 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C;Accession: F72709
R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Taka
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: F72709
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-146 <KAW>
A;Cross-references: DDBJ:AP000060; NID:g5104188; PIDN:BAA80078.1; PID:d1043864; PID:g51
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE1093

Query Match 75.7%; Score 28; DB 2; Length 146;
Best Local Similarity 66.7%; Pred. No. 66;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LAASLLSRV 9
|::|||:||
Db 74 LSSSLLTRV 82

RESULT 15
C97638
hypothetical protein AGR_C_4215 [imported] - Agrobacterium tumefaciens (strain C58, Cer
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
C;Accession: C97638
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tu
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: C97638
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-198 <KUR>
A;Cross-references: GB:AE007869; PIDN:AAK88060.1; PID:g15157484; GSPDB:GN00169
C;Genetics:
A;Gene: AGR_C_4215
A;Map position: circular chromosome

Query Match 75.7%; Score 28; DB 2; Length 198;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LAASLLS 7
|||||||
Db 14 LAASLLS 20

Search completed: October 5, 2004, 06:08:46
Job time : 15.1698 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 5, 2004, 06:07:39 ; Search time 5.09434 Seconds
(without alignments)
91.991 Million cell updates/sec

Title: US-10-022-286-1
Perfect score: 37
Sequence: 1 LAASLLSRV 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	30	81.1	422	1	MURA_HELPY	P56189 helicobacte
2	30	81.1	838	1	OS94_MOUSE	P48722 mus musculu
3	30	81.1	839	1	OS94_HUMAN	O95757 homo sapien
4	29	78.4	502	1	C72H_ARATH	Q9ltm6 arabidopsis
5	29	78.4	832	1	NAH3_RABIT	P26432 oryctolagus
6	29	78.4	865	1	FLUG_EMENI	P38094 emericella
7	29	78.4	886	1	HS97_STRFN	Q94738 strongyloce
8	29	78.4	889	1	HS97_STRPU	Q06068 strongyloce
9	29	78.4	1278	1	DHBF_BACSU	P45745 bacillus su
10	28	75.7	242	1	GLUA_CORGL	P48243 corynebacte
11	28	75.7	282	1	DRN1_CHICK	Q9ygi5 gallus gall
12	28	75.7	312	1	DHYS_SULSO	Q97zf1 sulfolobus
13	28	75.7	450	1	LAT_NOCLA	Q05174 nocardia la
14	28	75.7	493	1	GSPE_ECOLI	P45759 escherichia
15	28	75.7	497	1	GSPE_KLEPN	P15645 klebsiella
16	28	75.7	502	1	C72G_ARATH	Q9ltm7 arabidopsis
17	28	75.7	507	1	TYPH_RALSO	Q8y2x7 ralstonia s
18	28	75.7	531	1	FNX1_SCHPO	Q09752 schizosacch
19	28	75.7	738	1	TRFM_MOUSE	Q9r0r1 mus musculu
20	28	75.7	1468	1	CHD1_YEAST	P32657 saccharomyc
21	28	75.7	3144	1	VP13_YEAST	Q07878 saccharomyc
22	28	75.7	3214	1	BPA1_HUMAN	Q03001 homo sapien
23	27	73.0	136	1	AT91_MOUSE	P48202 mus musculu
24	27	73.0	136	1	AT91_RAT	Q06645 rattus norv
25	27	73.0	196	1	RAC2_LOTJA	Q40220 lotus japon
26	27	73.0	206	1	Y312_METJA	Q57760 methanococc
27	27	73.0	211	1	HS2M_SOYBN	Q39818 glycine max
28	27	73.0	234	1	RNH2_XYLFA	Q9pei7 xylella fas
29	27	73.0	234	1	RNH2_XYLFT	Q87ei6 xylella fas
30	27	73.0	244	1	ATPF_YEAST	P05626 saccharomyc
31	27	73.0	249	1	SURE_PSEPK	Q88mf1 pseudomonas
32	27	73.0	294	1	BLAN_SERMA	P52682 serratia ma
33	27	73.0	323	1	YVAX_BACSU	O34616 bacillus su
34	27	73.0	417	1	FEMB_STAEP	Q8cpa8 staphylococ
35	27	73.0	419	1	FEMB_STAAM	P14305 staphylococ
36	27	73.0	490	1	SUCP_LEUME	Q59495 leuconostoc
37	27	73.0	502	1	GSPE_PSEAE	Q00512 pseudomonas
38	27	73.0	516	1	BOCT_RAT	Q9p290 rattus norv
39	27	73.0	538	1	BOCT_HUMAN	Q8wug5 homo sapien
40	27	73.0	607	1	PEPF_MYCGE	P47429 mycoplasma
41	27	73.0	660	1	DNAK_CHLPN	P27542 chlamydia p
42	27	73.0	691	1	VP80_NPVAC	Q00733 autographa
43	27	73.0	1009	1	DPOL_HSVSA	P24907 herpesvirus
44	27	73.0	1100	1	JAK3_RAT	Q63272 rattus norv
45	27	73.0	1124	1	JAK3_HUMAN	P52333 homo sapien

ALIGNMENTS

```
RESULT 1
MURA_HELPY
ID   MURA_HELPY     STANDARD;      PRT;   422 AA.
AC   P56189;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   UDP-N-acetylglucosamine 1-carboxyvinyltransferase (EC 2.5.1.7)
DE   (Enoylpyruvate transferase) (UDP-N-acetylglucosamine enolpyruvyl
DE   transferase) (EPT).
GN   MURA OR MURZ OR HP0648.
OS   Helicobacter pylori (Campylobacter pylori).
OC   Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC   Helicobacteraceae; Helicobacter.
OX   NCBI_TaxID=210;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=26695 / ATCC 700392;
RX   MEDLINE=97394467; PubMed=9252185;
RA   Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA   Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA   Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA   Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA   McKenney K., FitzGerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA   Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA   Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA   Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA   Venter J.C.;
RT   "The complete genome sequence of the gastric pathogen Helicobacter
RT   pylori.";
RL   Nature 388:539-547(1997).
CC   -!- FUNCTION: Cell wall formation. Adds enolpyruvyl to UDP-N-
CC       acetylglucosamine (By similarity).
CC   -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + UDP-N-acetyl-D-
CC       glucosamine = phosphate + UDP-N-acetyl-3-(1-carboxyvinyl)-D-
CC       glucosamine.
CC   -!- PATHWAY: Peptidoglycan biosynthesis; first step.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: Belongs to the EPSP synthase family. MurA subfamily.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AE000578; AAD07708.1; -.
DR   PIR; H64600; H64600.
DR   HSSP; P33038; 1NAW.
DR   TIGR; HP0648; -.
DR   HAMAP; MF_00111; -; 1.
DR   InterPro; IPR005750; AcGlu_Tran_MurA.
DR   InterPro; IPR001986; EPSP_synth.
DR   Pfam; PF00275; EPSP_synthase; 1.
```

```
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   TIGRFAMs; TIGR01072; murA; 1.
KW   Peptidoglycan synthesis; Cell wall; Cell division; Transferase;
KW   Complete proteome.
FT   ACT_SITE    117    117       BINDS PEP (BY SIMILARITY).
SQ   SEQUENCE   422 AA;  45657 MW;  BC58A765A4F12362 CRC64;

  Query Match             81.1%;  Score 30;  DB 1;  Length 422;
  Best Local Similarity   87.5%;  Pred. No. 46;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LAASLLSR 8
              |||:||||
Db         29 LAATLLSR 36


RESULT 2
OS94_MOUSE
ID   OS94_MOUSE     STANDARD;      PRT;   838 AA.
AC   P48722; P97854; Q8BQD0; Q8CC45; Q91X29;
DT   01-FEB-1996 (Rel. 33, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Osmotic stress protein 94 (Heat shock 70-related protein APG-1).
GN   OSP94 OR APG1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1), AND TISSUE SPECIFICITY.
RX   MEDLINE=96218151; PubMed=8647834;
RA   Kojima R., Randall J., Brenner B.M., Gullans S.R.;
RT   "Osmotic stress protein 94 (Osp94). A new member of the Hsp110/SSE
RT   gene subfamily.";
RL   J. Biol. Chem. 271:12327-12332(1996).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 1), AND INDUCTION.
RC   STRAIN=DDY/STD; TISSUE=Testis;
RX   MEDLINE=97160564; PubMed=9006898;
RA   Kaneko Y., Nishiyama H., Nonoguchi K., Higashitsuji H., Kishishita M.,
RA   Fujita J.;
RT   "A novel hsp110-related gene, apg-1, that is abundantly expressed in
RT   the testis responds to a low temperature heat shock rather than the
RT   traditional elevated temperatures.";
RL   J. Biol. Chem. 272:2640-2645(1997).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   STRAIN=ddY; TISSUE=Testis;
RA   Kaneko Y., Fujita J.;
RT   "Apg-1b, an alternative form of apg-1 transcript.";
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   STRAIN=FVB/N; TISSUE=Brain, and Salivary gland;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [5]
RP   SEQUENCE OF 570-838 FROM N.A.
RX   MEDLINE=22354683; PubMed=12466851;
RA   Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA   Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA   Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA   Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA   Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA   Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA   Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA   Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA   Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA   Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA   Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA   Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA   Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA   Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA   Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA   Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA   Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA   Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA   Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA   Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA   Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA   Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA   Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA   Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA   Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA   Birney E., Hayashizaki Y.;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
CC   -!- FUNCTION: Possesses chaperone activity in vitro where it inhibits
CC       aggregation of citrate synthase (By similarity).
CC   -!- SUBUNIT: Homodimer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic. May translocate to the nucleus
CC       after heat shock (By similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=P48722-1; Sequence=Displayed;
CC       Name=2; Synonyms=Apg-1b;
CC         IsoId=P48722-2; Sequence=VSP_007500;
CC   -!- TISSUE SPECIFICITY: Highly expressed in testis. Also expressed in
CC       renal medulla of water-restricted animals.
CC   -!- INDUCTION: By hyperosmolar salt stress and heat shock.
CC   -!- SIMILARITY: Belongs to the heat shock protein 70 family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U23921; AAC52610.1; -.
DR   EMBL; D49482; BAA08446.1; -.
DR   EMBL; AB001926; BAA19468.1; -.
DR   EMBL; BC012712; AAH12712.1; -.
DR   EMBL; BC057002; AAH57002.1; -.
DR   EMBL; AK033950; BAC28524.1; -.
DR   EMBL; AK050997; BAC34491.1; -.
DR   MGD; MGI:107422; Osp94.
DR   GO; GO:0005737; C:cytoplasm; ISS.
DR   GO; GO:0005634; C:nucleus; ISS.
DR   GO; GO:0005524; F:ATP binding; ISS.
DR   GO; GO:0003773; F:heat shock protein activity; ISS.
DR   GO; GO:0006457; P:protein folding; ISS.
DR   InterPro; IPR001023; Hsp70.
```

```
DR   Pfam; PF00012; HSP70; 1.
DR   PRINTS; PR00301; HEATSHOCK70.
DR   ProDom; PD000089; Hsp70; 1.
DR   PROSITE; PS00297; HSP70_1; FALSE_NEG.
DR   PROSITE; PS00329; HSP70_2; 1.
DR   PROSITE; PS01036; HSP70_3; 1.
KW   Chaperone; Heat shock; ATP-binding; Alternative splicing.
FT   VARSPLIC      1     36       MSVVGIDLGFLNCYIAVARSGGIETIANEYSDRCTP -> M
FT                                GGPPRHGVLDREER (in isoform 2).
FT                                /FTId=VSP_007500.
FT   CONFLICT    175    176       TA -> HS (IN REF. 2 AND 3).
FT   CONFLICT    221    221       K -> E (IN REF. 2 AND 3).
FT   CONFLICT    279    279       A -> P (IN REF. 2 AND 3).
FT   CONFLICT    308    308       Q -> R (IN REF. 2 AND 3).
FT   CONFLICT    776    776       K -> M (IN REF. 1).
SQ   SEQUENCE   838 AA;  94381 MW;  B2C5847DA7EAF6B7 CRC64;

  Query Match             81.1%;  Score 30;  DB 1;  Length 838;
  Best Local Similarity   77.8%;  Pred. No. 92;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 LAASLLSRV 9
              | ||||:||
Db        309 LCASLLARV 317


RESULT 3
OS94_HUMAN
ID   OS94_HUMAN     STANDARD;      PRT;   839 AA.
AC   O95757; Q8IWA2;
DT   30-MAY-2000 (Rel. 39, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Osmotic stress protein 94 (Heat shock 70-related protein APG-1).
GN   OSP94 OR APG1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RX   MEDLINE=99453757; PubMed=10524232;
RA   Nonoguchi K., Itoh K., Xue J.H., Tokuchi H., Nishiyama H., Kaneko Y.,
RA   Tatsumi K., Okuno H., Tomiwa K., Fujita J.;
RT   "Cloning of human cDNAs for Apg-1 and Apg-2, members of the Hsp110
RT   family, and chromosomal assignment of their genes.";
RL   Gene 237:21-28(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: Possesses chaperone activity in vitro where it inhibits
CC       aggregation of citrate synthase (By similarity).
CC   -!- SUBUNIT: Homodimer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic. May translocate to the nucleus
CC       after heat shock (By similarity).
CC   -!- INDUCTION: By heat shock.
CC   -!- SIMILARITY: Belongs to the heat shock protein 70 family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB023421; BAA75063.1; -.
DR   EMBL; BC040560; AAH40560.1; -.
DR   GO; GO:0005737; C:cytoplasm; ISS.
DR   GO; GO:0005634; C:nucleus; ISS.
DR   GO; GO:0005524; F:ATP binding; ISS.
DR   GO; GO:0003773; F:heat shock protein activity; ISS.
DR   GO; GO:0006457; P:protein folding; ISS.
DR   InterPro; IPR001023; Hsp70.
DR   Pfam; PF00012; HSP70; 1.
DR   PRINTS; PR00301; HEATSHOCK70.
DR   ProDom; PD000089; Hsp70; 1.
DR   PROSITE; PS00297; HSP70_1; FALSE_NEG.
DR   PROSITE; PS00329; HSP70_2; 1.
DR   PROSITE; PS01036; HSP70_3; 1.
KW   Chaperone; Heat shock; ATP-binding.
FT   CONFLICT    211    211       S -> L (IN REF. 2).
FT   CONFLICT    806    806       H -> R (IN REF. 1).
SQ   SEQUENCE   839 AA;  94486 MW;  AB9936A3F32AAE8E CRC64;

  Query Match             81.1%;  Score 30;  DB 1;  Length 839;
  Best Local Similarity   77.8%;  Pred. No. 92;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 LAASLLSRV 9
              | ||||:||
Db        309 LCASLLARV 317


RESULT 4
C72H_ARATH
ID   C72H_ARATH     STANDARD;      PRT;   502 AA.
AC   Q9LTM6;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Cytochrome P450 71B17 (EC 1.14.-.-).
GN   CYP71B17 OR AT3G26160 OR MTC11.6.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=20277480; PubMed=10819329;
RA   Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Tabata S.;
RT   "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT   features of the regions of 4,504,864 bp covered by sixty P1 and TAC
RT   clones.";
RL   DNA Res. 7:131-135(2000).
CC   -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
```

```
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; AB024038; BAB02436.1; -.
DR   InterPro; IPR001128; Cytochrome_P450.
DR   Pfam; PF00067; p450; 1.
DR   PRINTS; PR00385; P450.
DR   PROSITE; PS00086; CYTOCHROME_P450; 1.
KW   Oxidoreductase; Monooxygenase; Transmembrane; Heme; Multigene family.
FT   TRANSMEM      1     21       POTENTIAL.
FT   METAL       444    444       IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ   SEQUENCE   502 AA;  57203 MW;  20B1B9E3AA702DB9 CRC64;

  Query Match             78.4%;  Score 29;  DB 1;  Length 502;
  Best Local Similarity   77.8%;  Pred. No. 92;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 LAASLLSRV 9
              ||||:| ||
Db        176 LAASILFRV 184


RESULT 5
NAH3_RABIT
ID   NAH3_RABIT     STANDARD;      PRT;   832 AA.
AC   P26432;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Sodium/hydrogen exchanger 3 (Na(+)/H(+) exchanger 3) (NHE-3).
GN   SLC9A3 OR NHE3.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=New Zealand white; TISSUE=Ileal villus, and Kidney cortex;
RX   MEDLINE=92250540; PubMed=1374392;
RA   Tse C.-M., Brant S.R., Walker S.S., Pouyssegur J., Donowitz M.;
RT   "Cloning and sequencing of a rabbit cDNA encoding an intestinal and
RT   kidney-specific Na+/H+ exchanger isoform (NHE-3).";
RL   J. Biol. Chem. 267:9340-9346(1992).
CC   -!- FUNCTION: Involved in pH regulation to eliminate acids generated
CC       by active metabolism or to counter adverse environmental
CC       conditions. Major proton extruding system driven by the inward
CC       sodium ion chemical gradient. Plays an important role in signal
CC       transduction.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Intestinal and kidney specific. Most abundant
CC       in kidney cortex, followed equally by ileum and ascending colon,
CC       then kidney medulla and jejunum. Is absent from duodenum and
CC       descending colon.
CC   -!- PTM: Phosphorylated (Possible).
CC   -!- SIMILARITY: Belongs to the Na(+)/H(+) exchanger family.
CC   -!- CAUTION: The number, localization and denomination of hydrophobic
CC       domains in the Na(+)/H(+) exchangers vary among authors.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; M87007; AAA31420.1; -.
DR   PIR; A40205; A40205.
DR   InterPro; IPR006153; Na_H_porter.
DR   InterPro; IPR004709; NaH_exchang.
DR   Pfam; PF00999; Na_H_Exchanger; 1.
DR   PRINTS; PR01084; NAHEXCHNGR.
DR   TIGRFAMs; TIGR00840; b_cpa1; 1.
KW   Transmembrane; Glycoprotein; Transport; Antiport; Sodium transport;
KW   Multigene family; Phosphorylation; Polymorphism.
FT   DOMAIN        1     11       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     12     27       M1 (POTENTIAL).
FT   DOMAIN       28     59       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     60     79       M2 (POTENTIAL).
FT   DOMAIN       80     81       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     82    101       M3 (POTENTIAL).
FT   DOMAIN      102    110       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    111    130       M4 (POTENTIAL).
FT   DOMAIN      131    134       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    135    154       M5 (POTENTIAL).
FT   DOMAIN      155    180       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    181    200       M5A (POTENTIAL).
FT   DOMAIN      201    209       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    210    229       M5B (POTENTIAL).
FT   DOMAIN      230    249       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    250    269       M6 (POTENTIAL).
FT   DOMAIN      270    298       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    299    319       M7 (POTENTIAL).
FT   DOMAIN      320    339       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    340    359       M8 (POTENTIAL).
FT   DOMAIN      360    366       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    367    385       M9 (POTENTIAL).
FT   DOMAIN      386    435       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    436    455       M10 (POTENTIAL).
FT   DOMAIN      456    832       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    325    325       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT     144    144       L -> P.
SQ   SEQUENCE   832 AA;  92748 MW;  8C8BB7C296CF8740 CRC64;

  Query Match             78.4%;  Score 29;  DB 1;  Length 832;
  Best Local Similarity   77.8%;  Pred. No. 1.5e+02;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 LAASLLSRV 9
              || |||:||
Db        766 LAPSLLARV 774


RESULT 6
FLUG_EMENI
ID   FLUG_EMENI     STANDARD;      PRT;   865 AA.
AC   P38094;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Protein fluG.
GN   FLUG OR ACOD.
OS   Emericella nidulans (Aspergillus nidulans).
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC   Eurotiales; Trichocomaceae; Emericella.
OX   NCBI_TaxID=162425;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FGSC 26;
RX   MEDLINE=95011568; PubMed=7926755;
RA   Lee B., Adams T.H.;
RT   "The Aspergillus nidulans fluG gene is required for production of an
RT   extracellular developmental signal and is related to prokaryotic
RT   glutamine synthetase I.";
RL   Genes Dev. 8:641-651(1994).
CC   -!- FUNCTION: May function as a GSI-related enzyme in synthesizing a
CC       small diffusible factor that acts as an extracellular signal
CC       directing asexual sporulation and perhaps other aspects of colony
CC       growth. May be involved in brlA activation (an early
CC       transcriptional regulator for conidiation specific gene).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: Belongs to the glutamine synthetase family.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
```

```
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L27817; AAC37414.1; -.
DR   PIR; A53186; A53186.
DR   InterPro; IPR008146; Gln_synt_C.
DR   Pfam; PF00120; gln-synt; 1.
DR   ProDom; PD001057; Gln_synt_C; 1.
KW   Conidiation.
FT   MUTAGEN     774    774       Y->N: TEMPERATURE-SENSITIVE.
SQ   SEQUENCE   865 AA;  96505 MW;  D17F3BB83B8F1719 CRC64;

  Query Match             78.4%;  Score 29;  DB 1;  Length 865;
  Best Local Similarity   66.7%;  Pred. No. 1.6e+02;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LAASLLSRV 9
              ||||:||::
Db        140 LAASVLSQI 148


RESULT 7
HS97_STRFN
ID   HS97_STRFN     STANDARD;      PRT;   886 AA.
AC   Q94738;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   97 kDa heat shock protein (Heat shock protein 110).
GN   HSP110.
OS   Strongylocentrotus franciscanus (Sea urchin).
OC   Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC   Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC   Strongylocentrotus.
OX   NCBI_TaxID=7665;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97287853; PubMed=9142981;
RA   Mauk R., Jaworski D., Kamei N., Glabe C.G.;
RT   "Identification of a 97-kDa heat shock protein from S. franciscanus
RT   ovaries with 94% amino acid identity to the S. purpuratus egg surface
RT   receptor for sperm.";
RL   Dev. Biol. 184:31-37(1997).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC   -!- SIMILARITY: Belongs to the heat shock protein 70 family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U69254; AAB09038.1; -.
DR   InterPro; IPR001023; Hsp70.
DR   Pfam; PF00012; HSP70; 1.
DR   PRINTS; PR00301; HEATSHOCK70.
DR   ProDom; PD000089; Hsp70; 3.
DR   PROSITE; PS00297; HSP70_1; FALSE_NEG.
DR   PROSITE; PS00329; HSP70_2; FALSE_NEG.
DR   PROSITE; PS01036; HSP70_3; 1.
KW   ATP-binding.
SQ   SEQUENCE   886 AA;  98446 MW;  252177643ECFEDD8 CRC64;

  Query Match             78.4%;  Score 29;  DB 1;  Length 886;
  Best Local Similarity   77.8%;  Pred. No. 1.6e+02;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 LAASLLSRV 9
              ||| || ||
Db        309 LAAELLKRV 317


RESULT 8
HS97_STRPU
ID   HS97_STRPU     STANDARD;      PRT;   889 AA.
AC   Q06068; Q94761;
DT   01-NOV-1997 (Rel. 35, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   97 kDa heat shock protein (Egg sperm receptor).
OS   Strongylocentrotus purpuratus (Purple sea urchin).
OC   Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC   Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC   Strongylocentrotus.
OX   NCBI_TaxID=7668;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 154-169 AND 481-495.
RC   TISSUE=Ovary;
RX   MEDLINE=93197888; PubMed=8383878;
RA   Foltz K.R., Partin J.S., Lennarz W.J.;
RT   "Sea urchin egg receptor for sperm: sequence similarity of binding
RT   domain and hsp70.";
RL   Science 259:1421-1425(1993).
RN   [2]
RP   REVISIONS.
RA   Lennarz W.J.;
RL   Submitted (SEP-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: CELL SURFACE RECOGNITION PROTEIN THAT BINDS ACROSOME-
CC       REACTED SPERM AND THEREBY MEDIATES BINDING AND SUBSEQUENT FUSION
CC       OF THE SPERM AND EGG.
CC   -!- SIMILARITY: Belongs to the heat shock protein 70 family.
CC   -!- CAUTION: WAS ORIGINALLY (REF.1) THOUGHT TO HAVE A N-TERMINAL
CC       SEQUENCE SIGNAL AND A C-TERMINAL TRANSMEMBRANE REGION. BOTH
CC       DOMAINS DO NOT EXIST IN THE REVISED SEQUENCE.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L04969; AAB09737.1; -.
DR   PIR; T11742; T11742.
DR   InterPro; IPR001023; Hsp70.
DR   Pfam; PF00012; HSP70; 1.
DR   PRINTS; PR00301; HEATSHOCK70.
DR   ProDom; PD000089; Hsp70; 2.
DR   PROSITE; PS00297; HSP70_1; FALSE_NEG.
DR   PROSITE; PS00329; HSP70_2; FALSE_NEG.
DR   PROSITE; PS01036; HSP70_3; 1.
KW   ATP-binding.
SQ   SEQUENCE   889 AA;  98618 MW;  1520EEDF70B0E0CF CRC64;

  Query Match             78.4%;  Score 29;  DB 1;  Length 889;
  Best Local Similarity   77.8%;  Pred. No. 1.6e+02;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 LAASLLSRV 9
              ||| || ||
Db        309 LAAELLKRV 317


RESULT 9
DHBF_BACSU
ID   DHBF_BACSU     STANDARD;      PRT;  1278 AA.
AC   P45745;
DT   01-NOV-1995 (Rel. 32, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Probable serine activating enzyme.
GN   DHBF OR BSU31960.
```

```
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=98044033; PubMed=9384377;
RA   Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA   Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA   Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA   Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA   Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA   Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA   Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA   Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA   Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA   Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA   Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA   Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA   Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA   Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA   Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA   Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA   Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA   Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA   Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA   Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA   Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA   Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA   Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA   Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA   Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA   Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
RA   Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA   Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT   "The complete genome sequence of the Gram-positive bacterium Bacillus
RT   subtilis.";
RL   Nature 390:249-256(1997).
RN   [2]
RP   SEQUENCE OF 1-95 FROM N.A.
RC   STRAIN=168 / Marburg;
RX   MEDLINE=96146538; PubMed=8550523;
RA   Rowland B.M., Taber H.W.;
RT   "Duplicate isochorismate synthase genes of Bacillus subtilis:
RT   regulation and involvement in the biosyntheses of menaquinone and
RT   2,3-dihydroxybenzoate.";
RL   J. Bacteriol. 178:854-861(1996).
CC   -!- COFACTOR: Contains 1 covalently bound phosphopantetheine (By
CC       similarity).
CC   -!- PATHWAY: 2,3-dihydroxybenzoate biosynthesis.
CC   -!- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
CC       family.
CC   -!- SIMILARITY: Contains 1 acyl carrier domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z99120; CAB15186.1; -.
DR   EMBL; U26444; AAC44634.1; -.
DR   PIR; E69615; E69615.
DR   HSSP; P14687; 1AMU.
DR   SubtiList; BG11243; dhbF.
DR   InterPro; IPR000873; AMP-bind.
DR   InterPro; IPR001242; Condensatn.
DR   InterPro; IPR006163; Pp_bind.
DR   InterPro; IPR006162; Ppantne_S.
DR   Pfam; PF00501; AMP-binding; 1.
DR   Pfam; PF00668; Condensation; 2.
DR   Pfam; PF00550; pp-binding; 1.
DR   PRINTS; PR00154; AMPBINDING.
DR   PROSITE; PS00012; PHOSPHOPANTETHEINE; FALSE_NEG.
DR   PROSITE; PS00455; AMP_BINDING; 1.
DR   PROSITE; PS50075; ACP_DOMAIN; 1.
KW   Ligase; Phosphopantetheine; Complete proteome.
FT   DOMAIN      966   1033       ACYL CARRIER (ACP).
FT   BINDING     996    996       PHOSPHOPANTETHEINE (POTENTIAL).
SQ   SEQUENCE   1278 AA;  142068 MW;  B44FBFF3FCB085B4 CRC64;

  Query Match             78.4%;  Score 29;  DB 1;  Length 1278;
  Best Local Similarity   66.7%;  Pred. No. 2.3e+02;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 LAASLLSRV 9
              ||| |:||:
Db        998 LAARLMSRI 1006


RESULT 10
GLUA_CORGL
ID   GLUA_CORGL     STANDARD;      PRT;   242 AA.
AC   P48243;
DT   01-FEB-1996 (Rel. 33, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Glutamate transport ATP-binding protein gluA.
GN   GLUA OR CGL1950.
OS   Corynebacterium glutamicum (Brevibacterium flavum).
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX   NCBI_TaxID=1718;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RX   MEDLINE=95173089; PubMed=7868586;
RA   Kronemeyer W., Peekhaus N., Kraemer R., Sahm H., Eggeling L.;
RT   "Structure of the gluABCD cluster encoding the glutamate uptake
RT   system of Corynebacterium glutamicum.";
RL   J. Bacteriol. 177:1152-1158(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA   Nakagawa S.;
RT   "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM
CC       FOR GLUTAMATE. PROBABLY RESPONSIBLE FOR ENERGY COUPLING TO THE
CC       TRANSPORT SYSTEM.
CC   -!- SUBCELLULAR LOCATION: Membrane-associated (Potential).
CC   -!- SIMILARITY: Belongs to the ABC transporter family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X81191; CAA57060.1; -.
DR   EMBL; AP005280; BAB99343.1; -.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR003439; ABC_transporter.
DR   Pfam; PF00005; ABC_tran; 1.
DR   ProDom; PD000006; ABC_transporter; 1.
DR   SMART; SM00382; AAA; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR   PROSITE; PS50893; ABC_TRANSPORTER_2; 1.
KW   Amino-acid transport; Transport; Membrane; ATP-binding;
KW   Complete proteome.
FT   NP_BIND      34     41       ATP (POTENTIAL).
FT   CONFLICT    149    149       R -> C (IN REF. 1).
```

```
SQ   SEQUENCE   242 AA;  26538 MW;  94401A263DE00194 CRC64;

  Query Match             75.7%;  Score 28;  DB 1;  Length 242;
  Best Local Similarity   77.8%;  Pred. No. 74;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 LAASLLSRV 9
              || ||| ||
Db        116 LAMSLLERV 124


RESULT 11
DRN1_CHICK
ID   DRN1_CHICK     STANDARD;      PRT;   282 AA.
AC   Q9YGI5;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Deoxyribonuclease I precursor (EC 3.1.21.1) (DNase I).
GN   DNASE1.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC   TISSUE=Pancreas;
RX   MEDLINE=22624822; PubMed=12739897;
RA   Hu C.C., Lu S.C., Cheng C.C., Chen L.H., Liao T.H.;
RT   "Chicken deoxyribonuclease: purification, characterization, gene
RT   cloning and gene expression.";
RL   J. Protein Chem. 22:41-49(2003).
CC   -!- FUNCTION: Among other functions, seems to be involved in cell
CC       death by apoptosis. Binds specifically to G-actin and blocks actin
CC       polymerization (By similarity).
CC   -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC       phosphodinucleotide and 5'-phosphooligonucleotide end-products.
CC   -!- COFACTOR: Divalent cations, particularly calcium and magnesium (By
CC       similarity).
CC   -!- SUBCELLULAR LOCATION: Secretory protein, stored in zymogen
CC       granules and found in the nuclear envelope (By similarity).
CC   -!- PTM: N-glycosylated.
CC   -!- SIMILARITY: Belongs to the DNase I family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ131751; CAA10503.1; -.
DR   HSSP; P00639; 1DNK.
DR   InterPro; IPR008185; DNase_I.
DR   InterPro; IPR008186; DNase_I_N.
DR   InterPro; IPR005135; Exo_endo_phos.
DR   Pfam; PF03372; Exo_endo_phos; 1.
DR   PRINTS; PR00130; DNASEI.
DR   ProDom; PD005408; DNase_I_N; 1.
DR   SMART; SM00476; DNaseIc; 1.
DR   PROSITE; PS00919; DNASE_I_1; FALSE_NEG.
DR   PROSITE; PS00918; DNASE_I_2; 1.
KW   Hydrolase; Endonuclease; Nuclease; Glycoprotein; Calcium; Signal;
KW   Apoptosis; Actin-binding.
FT   SIGNAL        1     20       POTENTIAL.
FT   CHAIN        21    282       DEOXYRIBONUCLEASE I.
FT   DISULFID    121    124       BY SIMILARITY.
FT   DISULFID    193    229       ESSENTIAL FOR ENZYMATIC ACTIVITY
FT                                (BY SIMILARITY).
FT   ACT_SITE     98     98       BY SIMILARITY.
FT   ACT_SITE    154    154       BY SIMILARITY.
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   282 AA;  31400 MW;  2226267860842EA7 CRC64;

  Query Match             75.7%;  Score 28;  DB 1;  Length 282;
  Best Local Similarity   77.8%;  Pred. No. 87;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 LAASLLSRV 9
              |||:|| ||
Db          9 LAAALLLRV 17


RESULT 12
DHYS_SULSO
ID   DHYS_SULSO     STANDARD;      PRT;   312 AA.
AC   Q97ZF1;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Probable deoxyhypusine synthase (EC 2.5.1.46) (DHS).
GN   DYS OR SSO0967.
OS   Sulfolobus solfataricus.
OC   Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC   Sulfolobus.
OX   NCBI_TaxID=2287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35092 / DSM 1617 / P2;
RX   MEDLINE=21332296; PubMed=11427726;
RA   She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA   Awayez M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,
RA   De Moors A., Erauso G., Fletcher C., Gordon P.M.K.,
RA   Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA   Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA   Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA   Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT   "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
CC   -!- FUNCTION: Catalyzes the NAD-dependent oxidative cleavage of
CC       spermidine and the subsequent transfer of the butylamine moiety of
CC       spermidine to the epsilon-amino group of a specific lysine residue
CC       of the eIF-5A precursor protein to form the intermediate
CC       deoxyhypusine residue (By similarity).
CC   -!- CATALYTIC ACTIVITY: [eIF5A-precursor]-lysine + spermidine =
CC       [eIF5A-precursor]-deoxyhypusine + propane-1,3-diamine.
CC   -!- COFACTOR: NAD (By similarity).
CC   -!- PATHWAY: Hypusine biosynthesis; first step.
CC   -!- SIMILARITY: Belongs to the deoxyhypusine synthase family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE006716; AAK41241.1; -.
DR   PIR; B90248; B90248.
DR   HAMAP; MF_00153; -; 1.
DR   InterPro; IPR002773; Deoxyhypus_synth.
DR   Pfam; PF01916; DS; 1.
DR   ProDom; PD007730; DS; 1.
KW   Hypusine biosynthesis; Transferase; NAD; Complete proteome.
SQ   SEQUENCE   312 AA;  35050 MW;  A9072CE18825A2D9 CRC64;

  Query Match             75.7%;  Score 28;  DB 1;  Length 312;
  Best Local Similarity   100.0%;  Pred. No. 96;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LAASLLS 7
              |||||||
Db        306 LAASLLS 312
```

```
RESULT 13
LAT_NOCLA
ID   LAT_NOCLA      STANDARD;      PRT;   450 AA.
AC   Q05174;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   L-lysine-epsilon aminotransferase (EC 2.6.1.36) (L-lysine
DE   aminotransferase) (Lysine 6-aminotransferase).
GN   LAT.
OS   Nocardia lactamdurans.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Pseudonocardineae; Pseudonocardiaceae; Amycolatopsis.
OX   NCBI_TaxID=1913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92011390; PubMed=1917857;
RA   Coque J., Liras P., Laiz L., Martin J.;
RT   "A gene encoding lysine 6-aminotransferase, which forms the
RT   beta-lactam precursor alpha-aminoadipic acid, is located in the
RT   cluster of cephamycin biosynthetic genes in Nocardia lactamdurans.";
RL   J. Bacteriol. 173:6258-6264(1991).
CC   -!- CATALYTIC ACTIVITY: L-lysine + 2-oxoglutarate = 2-aminoadipate 6-
CC       semialdehyde + L-glutamate.
CC   -!- COFACTOR: Pyridoxal phosphate.
CC   -!- PATHWAY: Cephamycin C biosynthesis; beta-lactam biosynthesis;
CC       first step.
CC   -!- SUBUNIT: Active as either a monomer or a homopolymer.
CC   -!- SIMILARITY: Belongs to class-III of pyridoxal-phosphate-dependent
CC       aminotransferases.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z21681; CAA79796.1; -.
DR   PIR; A38171; A38171.
DR   HSSP; P80147; 1GTX.
DR   InterPro; IPR005814; Aminotrans_3.
DR   Pfam; PF00202; aminotran_3; 1.
DR   PROSITE; PS00600; AA_TRANSFER_CLASS_3; 1.
KW   Transferase; Aminotransferase; Antibiotic biosynthesis;
KW   Pyridoxal phosphate.
FT   BINDING     300    300       PYRIDOXAL PHOSPHATE (POTENTIAL).
SQ   SEQUENCE   450 AA;  48806 MW;  C24915CAF2E93EE5 CRC64;

  Query Match             75.7%;  Score 28;  DB 1;  Length 450;
  Best Local Similarity   87.5%;  Pred. No. 1.4e+02;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 LAASLLSR 8
              ||||| ||
Db        442 LAASLASR 449


RESULT 14
GSPE_ECOLI
ID   GSPE_ECOLI     STANDARD;      PRT;   493 AA.
AC   P45759;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Probable general secretion pathway protein E (Type II traffic warden
DE   ATPase).
GN   GSPE OR B3326.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
CC   -!- FUNCTION: Involved in a general secretion pathway (GSP) for the
CC       export of proteins (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: BELONGS TO THE PULE/OUTE/EXEE/XPSE/XCPR FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U18997; AAA58123.1; -.
DR   EMBL; AE000409; AAC76351.1; -.
DR   PIR; A65126; A65126.
DR   EcoGene; EG12891; gspE.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR001482; GSPII_E.
DR   Pfam; PF00437; GSPII_E; 1.
DR   ProDom; PD000739; GSPII_E; 1.
DR   SMART; SM00382; AAA; 1.
DR   PROSITE; PS00662; T2SP_E; 1.
KW   Transport; ATP-binding; Complete proteome.
FT   NP_BIND     249    256       ATP (POTENTIAL).
SQ   SEQUENCE   493 AA;  54610 MW;  A13C67C3D48596CA CRC64;

  Query Match             75.7%;  Score 28;  DB 1;  Length 493;
  Best Local Similarity   66.7%;  Pred. No. 1.5e+02;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 LAASLLSRV 9
              ||| |:||:
Db        156 LAALLISRI 164


RESULT 15
GSPE_KLEPN
ID   GSPE_KLEPN     STANDARD;      PRT;   497 AA.
AC   P15645;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUL-1993 (Rel. 26, Last annotation update)
DE   General secretion pathway protein E (Type II traffic warden ATPase)
DE   (Pullulanase secretion protein pulE).
GN   PULE.
OS   Klebsiella pneumoniae.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Klebsiella.
OX   NCBI_TaxID=573;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92149318; PubMed=1738317;
RA   Pugsley A.P., D'Enfert C., Reyss I., Wandersman C.;
RT   "Pullulanase secretion in Escherichia coli K-12 requires a
RT   cytoplasmic protein and a putative polytopic cytoplasmic membrane
RT   protein.";
RL   Mol. Microbiol. 6:95-105(1992).
RN   [2]
RP   SEQUENCE OF 1-149 FROM N.A.
RX   MEDLINE=90008916; PubMed=2677007;
```

```
RA   D'Enfert C., Reyss I., Wandersman C., Pugsley A.P.;
RT   "Protein secretion by Gram-negative bacteria. Characterization of two
RT   membrane proteins required for pullulanase secretion by Escherichia
RT   coli K-12.";
RL   J. Biol. Chem. 264:17462-17468(1989).
CC   -!- FUNCTION: INVOLVED IN A GENERAL SECRETION PATHWAY (GSP) FOR THE
CC       EXPORT OF PROTEINS. REQUIRED FOR THE TRANSLOCATION OF PULLULANASE.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: BELONGS TO THE PULE/OUTE/EXEE/XPSE/XCPR FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M32613; AAA25127.1; -.
DR   PIR; S20034; C34469.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR001482; GSPII_E.
DR   Pfam; PF00437; GSPII_E; 1.
DR   ProDom; PD000739; GSPII_E; 1.
DR   SMART; SM00382; AAA; 1.
DR   PROSITE; PS00662; T2SP_E; 1.
KW   Transport; ATP-binding.
FT   NP_BIND     255    262       ATP (POTENTIAL).
SQ   SEQUENCE   497 AA;  55129 MW;  7486926B51728D3B CRC64;

  Query Match             75.7%;  Score 28;  DB 1;  Length 497;
  Best Local Similarity   66.7%;  Pred. No. 1.5e+02;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 LAASLLSRV 9
              ||| |:||:
Db        162 LAALLISRI 170


Search completed: October  5, 2004, 06:17:25
Job time : 10.0943 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 5, 2004, 06:07:39 ; Search time 26.8302 Seconds
(without alignments)
105.838 Million cell updates/sec

Title: US-10-022-286-1
Perfect score: 37
Sequence: 1 LAASLLSRV 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	37	100.0	637	16	O53585	O53585 mycobacteri
2	37	100.0	637	16	Q7TVM4	Q7tvm4 mycobacteri
3	34	91.9	501	5	Q9NE29	Q9ne29 leishmania
4	33	89.2	1326	16	Q8P449	Q8p449 xanthomonas
5	32	86.5	800	11	Q8R3J3	Q8r3j3 mus musculu
6	32	86.5	1062	11	Q8BSZ0	Q8bsz0 mus musculu
7	31	83.8	230	13	Q7ZVA1	Q7zva1 brachydanio
8	31	83.8	233	10	Q8SAZ2	Q8saz2 oryza sativ
9	31	83.8	411	16	Q82Q69	Q82q69 streptomyce
10	31	83.8	559	2	Q9S3V6	Q9s3v6 streptomyce
11	31	83.8	677	2	Q9F2E7	Q9f2e7 streptomyce
12	31	83.8	677	2	Q84DH9	Q84dh9 streptomyce
13	31	83.8	1777	12	O89278	O89278 himetobi p
14	31	83.8	2362	13	Q7T160	Q7t160 brachydanio
15	31	83.8	3604	5	Q9VYK0	Q9vyk0 drosophila
16	30	81.1	258	5	Q9GRQ9	Q9grq9 leishmania
17	30	81.1	267	10	Q9LSU9	Q9lsu9 arabidopsis
18	30	81.1	290	16	P71896	P71896 mycobacteri
19	30	81.1	290	16	Q7TYU3	Q7tyu3 mycobacteri
20	30	81.1	296	2	O33353	O33353 mycobacteri
21	30	81.1	458	17	Q8TVI3	Q8tvi3 methanopyru
22	30	81.1	489	10	Q8W0J7	Q8w0j7 oryza sativ
23	30	81.1	827	4	Q9NS96	Q9ns96 homo sapien
24	30	81.1	851	5	Q8IKY9	Q8iky9 plasmodium
25	30	81.1	1086	4	Q8N4T5	Q8n4t5 homo sapien
26	30	81.1	1363	5	Q9NEI0	Q9nei0 leishmania
27	30	81.1	1564	13	Q8QG98	Q8qg98 raja erinac
28	30	81.1	2159	11	Q91V24	Q91v24 mus musculu
29	30	81.1	3073	2	Q8KUH4	Q8kuh4 actinosynne
30	30	81.1	9376	2	O85168	O85168 pseudomonas
31	29	78.4	120	16	Q8PFF0	Q8pff0 xanthomonas
32	29	78.4	151	5	O18573	O18573 andrena eri
33	29	78.4	186	5	O44159	O44159 caenorhabdi
34	29	78.4	255	16	Q82MJ2	Q82mj2 streptomyce
35	29	78.4	265	16	Q8G5Y7	Q8g5y7 bifidobacte
36	29	78.4	280	17	Q8PUG0	Q8pug0 methanosarc
37	29	78.4	374	17	Q97BE7	Q97be7 thermoplasm
38	29	78.4	386	16	Q82D64	Q82d64 streptomyce
39	29	78.4	402	17	Q8ZSX9	Q8zsx9 pyrobaculum
40	29	78.4	419	2	Q9X9D7	Q9x9d7 staphylococ
41	29	78.4	438	16	Q9X7U6	Q9x7u6 streptomyce
42	29	78.4	505	10	Q9XJ15	Q9xj15 oryza sativ
43	29	78.4	564	9	Q9G0G1	Q9g0g1 roseophage
44	29	78.4	568	16	Q884D6	Q884d6 pseudomonas
45	29	78.4	715	4	Q8N590	Q8n590 homo sapien

ALIGNMENTS

RESULT 1
O53585
ID O53585 PRELIMINARY; PRT; 637 AA.
AC O53585;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein Rv3808c.
GN RV3808C OR MT3915 OR MTV026.13C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.

```
DR   EMBL; AL022076; CAA17872.1; -.
DR   EMBL; AE007185; AAK48281.1; -.
DR   PIR; D70888; D70888.
DR   TIGR; MT3915; -.
DR   TubercuList; Rv3808c; -.
DR   InterPro; IPR001173; Glyco_trans_2.
KW   Hypothetical protein; Complete proteome.
FT   CONFLICT    482    482       Q -> R (IN REF. 2).
SQ   SEQUENCE   637 AA;  71506 MW;  D7C9AB28B5005B66 CRC64;

  Query Match             100.0%;  Score 37;  DB 16;  Length 637;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LAASLLSRV 9
              |||||||||
Db          4 LAASLLSRV 12


RESULT 2
Q7TVM4
ID   Q7TVM4      PRELIMINARY;      PRT;   637 AA.
AC   Q7TVM4;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Galactofuranosyl transferase (EC 2.-).
GN   MB3838C.
OS   Mycobacterium bovis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1765;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AF2122/97;
RX   MEDLINE=22709107; PubMed=12788972;
RA   Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA   Pryor M., Duthoy S., Grondin S., Lacroix C., Monsempe C., Simon S.,
RA   Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA   Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT   "The complete genome sequence of Mycobacterium bovis.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR   EMBL; BX248347; CAD96024.1; -.
KW   Transferase; Complete proteome.
SQ   SEQUENCE   637 AA;  71506 MW;  D7C9AB28B5005B66 CRC64;

  Query Match             100.0%;  Score 37;  DB 16;  Length 637;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LAASLLSRV 9
              |||||||||
Db          4 LAASLLSRV 12


RESULT 3
Q9NE29
ID   Q9NE29      PRELIMINARY;      PRT;   501 AA.
AC   Q9NE29;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Hypothetical predicted protein L3204.05, unknown function.
GN   L3204.05.
OS   Leishmania major.
OC   Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX   NCBI_TaxID=5664;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Friedlin;
RA   Fuchs M., Gabel C., Mueller-Auer S., Schaefer M., Rieger M.,
RA   Ivens A.C., Quail M., Rajandream M.A., Barrell B.G.;
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Friedlin;
RX   MEDLINE=98146435; PubMed=9477341;
RA   Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RA   Smith D.F.;
RT   "A physical map of the Leishmania major Friedlin genome.";
RL   Genome Res. 8:135-145(1998).
DR   EMBL; AL354552; CAB89627.1; -.
SQ   SEQUENCE   501 AA;  56187 MW;  052FFAF93DAAE00B CRC64;

  Query Match             91.9%;  Score 34;  DB 5;  Length 501;
  Best Local Similarity   88.9%;  Pred. No. 47;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LAASLLSRV 9
              |:|||||||
Db         16 LSASLLSRV 24


RESULT 4
Q8P449
ID   Q8P449      PRELIMINARY;      PRT;  1326 AA.
AC   Q8P449;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   ATP-dependent serine activating enzyme.
GN   ENTF OR XCC3867.
OS   Xanthomonas campestris (pv. campestris).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC   Xanthomonadaceae; Xanthomonas.
OX   NCBI_TaxID=340;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 33913 / NCPPB 528;
RX   MEDLINE=22022145; PubMed=12024217;
RA   da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA   Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA   Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA   Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA   Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA   Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA   Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA   Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA   Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA   Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA   Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA   Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA   Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA   Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA   Setubal J.C., Kitajima J.P.;
RT   "Comparison of the genomes of two Xanthomonas pathogens with differing
RT   host specificities.";
RL   Nature 417:459-463(2002).
DR   EMBL; AE012507; AAM43098.1; -.
DR   GO; GO:0003824; F:catalytic activity; IEA.
DR   GO; GO:0016788; F:hydrolase activity, acting on ester bonds; IEA.
DR   GO; GO:0009058; P:biosynthesis; IEA.
DR   InterPro; IPR000873; AMP-bind.
DR   InterPro; IPR001242; Condensatn.
DR   InterPro; IPR006162; Ppantne_S.
DR   InterPro; IPR006163; Pp_bind.
DR   InterPro; IPR001031; Thioesterase.
DR   Pfam; PF00501; AMP-binding; 1.
DR   Pfam; PF00668; Condensation; 1.
DR   Pfam; PF00550; pp-binding; 1.
DR   Pfam; PF00975; Thioesterase; 1.
DR   PROSITE; PS50075; ACP_DOMAIN; 1.
DR   PROSITE; PS00455; AMP_BINDING; 1.
DR   PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.
KW   Complete proteome.
SQ   SEQUENCE   1326 AA;  145209 MW;  4071EBF84D636932 CRC64;
```

```
  Query Match             89.2%;  Score 33;  DB 16;  Length 1326;
  Best Local Similarity   88.9%;  Pred. No. 2.1e+02;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 LAASLLSRV 9
              ||| |||||
Db        724 LAAQLLSRV 732


RESULT 5
Q8R3J3
ID   Q8R3J3      PRELIMINARY;      PRT;   800 AA.
AC   Q8R3J3;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC025182; AAH25182.1; -.
KW   Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   800 AA;  89878 MW;  F88849947E435D26 CRC64;

  Query Match             86.5%;  Score 32;  DB 11;  Length 800;
  Best Local Similarity   88.9%;  Pred. No. 2.1e+02;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 LAASLLSRV 9
              |||||| ||
Db        100 LAASLLHRV 108


RESULT 6
Q8BSZ0
ID   Q8BSZ0      PRELIMINARY;      PRT;  1062 AA.
AC   Q8BSZ0;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical P-loop containing nucleotide triphosphate hydrolases
DE   structure containing protein.
GN   E130315B21RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Placenta, and Extraembryonic tissue;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK028381; BAC25919.1; -.
DR   PIR; PT0551; PT0551.
DR   PIR; PT0597; PT0597.
DR   MGD; MGI:2443732; E130315B21Rik.
KW   Hypothetical protein.
SQ   SEQUENCE   1062 AA;  119461 MW;  3502AD466A2806FD CRC64;

  Query Match             86.5%;  Score 32;  DB 11;  Length 1062;
  Best Local Similarity   88.9%;  Pred. No. 2.8e+02;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 LAASLLSRV 9
              |||||| ||
Db        362 LAASLLHRV 370


RESULT 7
Q7ZVA1
ID   Q7ZVA1      PRELIMINARY;      PRT;   230 AA.
AC   Q7ZVA1;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical protein (Fragment).
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Body;
RA   Strausberg R.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC045944; AAH45944.1; -.
KW   Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   230 AA;  26907 MW;  17E478CB58DE1502 CRC64;

  Query Match             83.8%;  Score 31;  DB 13;  Length 230;
  Best Local Similarity   87.5%;  Pred. No. 1e+02;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LAASLLSR 8
              :|||||||
Db         10 MAASLLSR 17


RESULT 8
Q8SAZ2
ID   Q8SAZ2      PRELIMINARY;      PRT;   233 AA.
AC   Q8SAZ2;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Putative dimethyladenosine transferse.
GN   OSJNBA0029P16.19 OR OSJNAB0022I16.13.
OS   Oryza sativa (Rice), and
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530, 39947;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   SPECIES=O.sativa;
RA   Wing R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,
RA   Saski C., Henry D., Oates R., Simmons J., Wilson R., Minx P., Du H.;
RT   "Rice Genomic Sequence.";
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   SPECIES=O.sativa (japonica cultivar-group);
RA   Wing R.A., Yu Y., Yang T.J., Nah G., Soderlund C., Chen M., Kim H.-R.,
RA   Rambo T., Saski C., Henry D., Oates R., Simmons J.;
RT   "Rice Genomic Sequence.";
RL   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   SPECIES=O.sativa (japonica cultivar-group); STRAIN=cv. Nipponbare;
RA   The Rice Chromosome 10 Sequencing Consortium;
RT   "In-depth view of structure, activity, and evolution of rice
RT   chromosome 10.";
```

```
RL   Science 300:1566-1569(2003).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   SPECIES=O.sativa (japonica cultivar-group); STRAIN=cv. Nipponbare;
RA   Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
RL   Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC093568; AAL78109.1; -.
DR   EMBL; AC122146; AAM47293.1; -.
DR   EMBL; AE017065; AAP52468.1; -.
DR   Gramene; Q8SAZ2; -.
DR   GO; GO:0000179; F:rRNA (adenine-N6,N6-)-dimethyltransferase a. . .; IEA.
DR   GO; GO:0008649; F:rRNA methyltransferase activity; IEA.
DR   GO; GO:0000154; P:rRNA modification; IEA.
DR   InterPro; IPR001737; RRNA_A_dimeth.
DR   Pfam; PF00398; RrnaAD; 1.
SQ   SEQUENCE   233 AA;  26700 MW;  173EF317C1E8B13C CRC64;

  Query Match             83.8%;  Score 31;  DB 10;  Length 233;
  Best Local Similarity   77.8%;  Pred. No. 1e+02;
  Matches    7;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LAASLLSRV 9
              |::||||||
Db         51 LSSSLLSRV 59


RESULT 9
Q82Q69
ID   Q82Q69      PRELIMINARY;      PRT;   411 AA.
AC   Q82Q69;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Putative fosmidmycin resistance protein.
GN   SAV652.
OS   Streptomyces avermitilis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=33903;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX   MEDLINE=21477403; PubMed=11572948;
RA   Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA   Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA   Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT   "Genome sequence of an industrial microorganism Streptomyces
RT   avermitilis: deducing the ability of producing secondary
RT   metabolites.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX   MEDLINE=22608306; PubMed=12692562;
RA   Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA   Sakaki Y., Hattori M., Omura S.;
RT   "Complete genome sequence and comparative analysis of the industrial
RT   microorganism Streptomyces avermitilis.";
RL   Nat. Biotechnol. 21:526-531(2003).
DR   EMBL; AP005023; BAC68362.1; -.
DR   InterPro; IPR007114; MFS.
DR   PROSITE; PS50850; MFS; 1.
KW   Complete proteome.
SQ   SEQUENCE   411 AA;  41906 MW;  5ADA60E6719DBF71 CRC64;

  Query Match             83.8%;  Score 31;  DB 16;  Length 411;
  Best Local Similarity   88.9%;  Pred. No. 1.8e+02;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 LAASLLSRV 9
              |||||||.|
Db         47 LAASLLSSV 55


RESULT 10
Q9S3V6
ID   Q9S3V6      PRELIMINARY;      PRT;   559 AA.
AC   Q9S3V6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Tendae peptide synthetase A (Fragment).
GN   TPSA.
OS   Streptomyces tendae.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 31160;
RA   Engel P., Scharfenstein L.L.;
RT   "Isolation of a gene specifying a peptide synthetase protein from
RT   Streptomyces tendae.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF172065; AAD51026.1; -.
DR   HSSP; P14687; 1AMU.
DR   GO; GO:0003824; F:catalytic activity; IEA.
DR   GO; GO:0008152; P:metabolism; IEA.
DR   InterPro; IPR000873; AMP-bind.
DR   InterPro; IPR006163; Pp_bind.
DR   Pfam; PF00501; AMP-binding; 1.
DR   Pfam; PF00550; pp-binding; 1.
DR   PRINTS; PR00154; AMPBINDING.
DR   PROSITE; PS50075; ACP_DOMAIN; 1.
DR   PROSITE; PS00455; AMP_BINDING; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   559 AA;  60057 MW;  FED08245BFA0647F CRC64;

  Query Match             83.8%;  Score 31;  DB 2;  Length 559;
  Best Local Similarity   66.7%;  Pred. No. 2.5e+02;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LAASLLSRV 9
              |||||::|:
Db        510 LAASLIARI 518


RESULT 11
Q9F2E7
ID   Q9F2E7      PRELIMINARY;      PRT;   677 AA.
AC   Q9F2E7;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   NikP1 protein.
GN   NIKP1.
OS   Streptomyces tendae.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Tue901;
RX   MEDLINE=21080548; PubMed=11212921;
RA   Lauer B., Russwurm R., Schwarz W., Kalmanchelyi A., Bruntner C.,
RA   Rosemeier A., Bormann C.;
RT   "Molecular characterization of co-transcribed genes from Streptomyces
RT   tendae Tue901 involved in the biosynthesis of the peptidyl moiety and
RT   assembly of the peptidyl nucleoside antibiotic nikkomycin.";
RL   Mol. Gen. Genet. 264:662-673(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Tue901;
RX   MEDLINE=20177695; PubMed=10712601;
RA   Lauer B., Russwurm R., Bormann C.;
RT   "Molecular characterization of two genes from Streptomyces tendae
```

```
RT   Tu901 required for the formation of the 4-formyl-4-imidazolin-2-one
RT   containing nucleoside moiety of the peptidyl nucleoside antibiotic
RT   nikkomycin.";
RL   Eur. J. Biochem. 267:1698-1706(2000).
DR   EMBL; AJ250581; CAC11137.1; -.
DR   HSSP; P14687; 1AMU.
DR   GO; GO:0003824; F:catalytic activity; IEA.
DR   GO; GO:0008152; P:metabolism; IEA.
DR   InterPro; IPR000873; AMP-bind.
DR   InterPro; IPR005153; MbtH.
DR   InterPro; IPR006163; Pp_bind.
DR   Pfam; PF00501; AMP-binding; 1.
DR   Pfam; PF03621; MbtH; 1.
DR   Pfam; PF00550; pp-binding; 1.
DR   PRINTS; PR00154; AMPBINDING.
DR   PROSITE; PS50075; ACP_DOMAIN; 1.
DR   PROSITE; PS00455; AMP_BINDING; 1.
KW   Phosphopantetheine.
SQ   SEQUENCE   677 AA;  73093 MW;  558313356C5F5581 CRC64;

  Query Match             83.8%;  Score 31;  DB 2;  Length 677;
  Best Local Similarity   66.7%;  Pred. No. 3e+02;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LAASLLSRV 9
              |||||::|:
Db        628 LAASLIARI 636


RESULT 12
Q84DH9
ID   Q84DH9      PRELIMINARY;      PRT;   677 AA.
AC   Q84DH9;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Peptide synthase.
GN   SANO.
OS   Streptomyces ansochromogenes.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=115647;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Wang G., Nie L., Tan H.;
RT   "Cloning and characterization of sanO, a gene involoved in nikkomycin
RT   biosynthesis of Streptomyces ansochromogenes.";
RL   Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY188796; AAO73548.1; -.
DR   GO; GO:0003824; F:catalytic activity; IEA.
DR   GO; GO:0008152; P:metabolism; IEA.
DR   InterPro; IPR000873; AMP-bind.
DR   InterPro; IPR005153; MbtH.
DR   InterPro; IPR006163; Pp_bind.
DR   Pfam; PF00501; AMP-binding; 1.
DR   Pfam; PF03621; MbtH; 1.
DR   Pfam; PF00550; pp-binding; 1.
DR   PRINTS; PR00154; AMPBINDING.
DR   PROSITE; PS50075; ACP_DOMAIN; 1.
DR   PROSITE; PS00455; AMP_BINDING; 1.
SQ   SEQUENCE   677 AA;  73000 MW;  2C98807709862F8C CRC64;

  Query Match             83.8%;  Score 31;  DB 2;  Length 677;
  Best Local Similarity   66.7%;  Pred. No. 3e+02;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LAASLLSRV 9
              |||||::|:
Db        628 LAASLIARI 636


RESULT 13
O89278
ID   O89278      PRELIMINARY;      PRT;  1777 AA.
AC   O89278;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Nonstructural protein.
OS   Himetobi P virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Dicistroviridae;
OC   Cripavirus.
OX   NCBI_TaxID=81583;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20019726; PubMed=10550677;
RA   Nakashima N., Sasaki J., Toriyama S.;
RT   "Determining the nucleotide sequence and capsid-coding region of
RT   Himetobi P virus: a member of a novel group of RNA viruses that infect
RT   insects.";
RL   Arch. Virol. 144:2051-2058(1999).
DR   EMBL; AB017037; BAA32553.1; -.
DR   PIR; T00490; T00490.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0003724; F:RNA helicase activity; IEA.
DR   GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR   GO; GO:0006350; P:transcription; IEA.
DR   GO; GO:0019079; P:viral genome replication; IEA.
DR   InterPro; IPR004004; Calici_pol_hel.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR000605; RNA_helicase.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR001205; RNA_pol_P3D.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR   Pfam; PF00910; RNA_helicase; 1.
DR   PRINTS; PR00918; CALICVIRUSNS.
SQ   SEQUENCE   1777 AA;  201730 MW;  3E8F872B3F406E67 CRC64;

  Query Match             83.8%;  Score 31;  DB 12;  Length 1777;
  Best Local Similarity   66.7%;  Pred. No. 7.9e+02;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LAASLLSRV 9
              ||||:||::
Db        587 LAASILSKI 595


RESULT 14
Q7T160
ID   Q7T160      PRELIMINARY;      PRT;  2362 AA.
AC   Q7T160;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   SI:dZ146N9.1 (Novel protein similar to human general control of
DE   amino-acid synthesis 1-like 1 (Yeast) (GCN1)) (Fragment).
GN   SI:DZ146N9.1.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Beasley H.;
RL   Submitted (JUL-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL954693; CAE17591.1; -.
KW   Oncogene.
FT   NON_TER       1      1
SQ   SEQUENCE   2362 AA;  259290 MW;  498508A0EE8B4579 CRC64;

  Query Match             83.8%;  Score 31;  DB 13;  Length 2362;
  Best Local Similarity   77.8%;  Pred. No. 1.1e+03;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy          1 LAASLLSRV 9
              |||||: ||
Db        612 LAASLIKRV 620
```

RESULT 15
Q9VYK0
ID Q9VYK0 PRELIMINARY; PRT; 3604 AA.
AC Q9VYK0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CG4013 protein.
GN SMR OR CG4013.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AE003490; AAF48195.2; -.
DR FlyBase; FBgn0024308; Smr.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002086; Aldehyde_dehydr.
DR InterPro; IPR001005; Myb_DNA_binding.
DR Pfam; PF00249; myb_DNA-binding; 1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; 1.
SQ SEQUENCE 3604 AA; 378155 MW; B7563A180C1D546B CRC64;

Query Match 83.8%; Score 31; DB 5; Length 3604;
Best Local Similarity 66.7%; Pred. No. 1.6e+03;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 LAASLLSRV 9
              |:|||:||:
Db        122 LSASLISRI 130
```

Search completed: October 5, 2004, 06:16:53
Job time : 34.8302 secs

OM protein - protein search, using sw model

Run on: October 5, 2004, 06:07:39 ; Search time 40.9245 Seconds
(without alignments)
62.137 Million cell updates/sec

Title: US-10-022-286-1
Perfect score: 37
Sequence: 1 LAASLLSRV 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	37	100.0	9	5	ABG31245	Abg31245 Immunogen
2	37	100.0	637	5	ABU05979	Abu05979 M. tuberc
3	32	86.5	175	4	AAU65823	Aau65823 Propionib
4	32	86.5	175	6	ABM62342	Abm62342 Propionib
5	32	86.5	580	4	AAU48527	Aau48527 Propionib
6	32	86.5	580	6	ABM45046	Abm45046 Propionib
7	31	83.8	382	3	AAY82727	Aay82727 Thiosulph
8	31	83.8	386	3	AAY82725	Aay82725 Thiosulph
9	30	81.1	122	5	ABP35331	Abp35331 Human ORF
10	30	81.1	290	6	ABM15857	Abm15857 Mycobacte
11	30	81.1	422	4	AAU35758	Aau35758 Helicobac
12	30	81.1	422	6	ABU30835	Abu30835 Protein e
13	30	81.1	509	4	AAU62933	Aau62933 Propionib
14	30	81.1	509	4	ABG05437	Abg05437 Novel hum
15	30	81.1	509	4	ABG04461	Abg04461 Novel hum
16	30	81.1	509	6	ABM59452	Abm59452 Propionib
17	30	81.1	538	6	ABU33928	Abu33928 Protein e
18	30	81.1	572	4	AAB93547	Aab93547 Human pro
19	30	81.1	614	6	ABU32206	Abu32206 Protein e
20	30	81.1	707	6	ABM65866	Abm65866 Propionib
21	30	81.1	827	6	ABP98316	Abp98316 Amino aci
22	30	81.1	838	5	ABB57153	Abb57153 Mouse isc
23	30	81.1	1051	5	AAO22157	Aao22157 Ramoplani
24	30	81.1	1093	6	AAE34868	Aae34868 Human kin
25	30	81.1	1152	4	AAM40940	Aam40940 Human pol
26	30	81.1	2167	6	ABG72695	Abg72695 Mouse ATP
27	29	78.4	119	4	ABG26252	Abg26252 Novel hum
28	29	78.4	173	4	AAG81988	Aag81988 S. epider
29	29	78.4	173	4	AAG82804	Aag82804 S. epider
30	29	78.4	265	5	ABP65754	Abp65754 Bifidobac
31	29	78.4	393	6	ABU43855	Abu43855 Protein e
32	29	78.4	438	3	AAG50386	Aag50386 Arabidops
33	29	78.4	502	3	AAG50385	Aag50385 Arabidops
34	29	78.4	608	6	ABJ19373	Abj19373 NOVX rela
35	29	78.4	715	4	AAM78647	Aam78647 Human pro
36	29	78.4	715	4	ABU52753	Abu52753 Human nuc
37	29	78.4	760	4	ABG21925	Abg21925 Novel hum
38	29	78.4	815	4	ABG19101	Abg19101 Novel hum
39	29	78.4	1547	4	ABB66152	Abb66152 Drosophil
40	29	78.4	1591	4	ABB62158	Abb62158 Drosophil
41	29	78.4	2175	4	ABB59839	Abb59839 Drosophil
42	28	75.7	43	5	AAU77142	Aau77142 Human SCY
43	28	75.7	67	5	ABP05466	Abp05466 Human ORF
44	28	75.7	68	4	ABB15133	Abb15133 Human ner
45	28	75.7	70	3	AAB39117	Aab39117 Human sec

ALIGNMENTS

RESULT 1
ABG31245
ID ABG31245 standard; peptide; 9 AA.
XX
AC ABG31245;
XX
DT 05-NOV-2002 (first entry)
XX
DE Immunogenic peptide, #1, of M. tuberculosis hypothetical protein Rv3808c.
XX
KW Immunogen; tuberculosis; TB; tuberculostatic; HLA-A2;
KW human leukocyte antigen A2; major histocompatibility complex; MHC;
KW cytotoxic T lymphocyte; CTL; cytokine; immunotherapy; epitope;
KW therapeutic; MHC tetramer; lymphocyte; hypothetical protein; Rv3808c.
XX
OS Mycobacterium tuberculosis.
XX
PN WO200248175-A2.
XX
PD 20-JUN-2002.
XX
PF 12-DEC-2001; 2001WO-US048742.
XX
PR 13-DEC-2000; 2000US-0255292P.
PR 30-JAN-2001; 2001US-0264978P.
XX
PA (ARGO-) ARGONEX PHARM.
PA (UYVI-) UNIV VIRGINIA PATENT FOUND.
XX
PI Flyer D, Ross MM, Hunt DF, White FM;
XX
DR WPI; 2002-599500/64.
XX
PT Novel immunogen comprising a peptide segment derived from Mycobacterium
PT tuberculosis, useful for inducing a cytotoxic T lymphocyte response in
PT vivo to tuberculosis infected cells.
XX
PS Claim 2; Page 45; 52pp; English.
XX
CC The invention discloses an immunogen, and polynucleotide encoding it,
CC comprising a peptide segment derived from Mycobacterium tuberculosis,
CC where the immunogen in not heat shock protein (hsp) 65 protein. The
CC peptide, and antibody specific for the peptide, can be used to treat a
CC subject with tuberculosis (TB) characterised by tuberculosis infected
CC cells expressing human leukocyte antigen A2 (HLA-A2) or any class I major
CC histocompatibility complex (MHC) molecule. The treatment involves
CC administering cytotoxic T lymphocytes (CTLs), induced in vitro using the
CC peptide, to destroy infected cells through direct lysis or to effect the

CC destruction of the infected cells indirectly through the elaboration of
CC cytokines. The ability of the immunogens to generate CTLs in vitro can
CC serve as a diagnostic for tuberculosis and they can be useful for
CC stimulating production of antibodies for use in passive immunotherapy.
CC The peptides are useful for screening a sample for the presence of CTLs
CC that specifically recognise the corresponding epitopes, as a diagnostic
CC tool to evaluate the therapeutic efficacy of the immunotherapeutic
CC treatments and to prepare class I MHC tetramers which can be used in
CC conjunction with flow cytometry to quantitate the frequency of peptide-
CC specific CTLs that are present in a sample of lymphocytes from an
CC individual. The sequence presented is the immunogenic peptide, #1, of the
CC M. tuberculosis hypothetical protein Rv3808c
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 37; DB 5; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LAASLLSRV 9
|||||||||
Db 1 LAASLLSRV 9

RESULT 2
ABU05979
ID ABU05979 standard; protein; 637 AA.
XX
AC ABU05979;
XX
DT 08-APR-2003 (first entry)
XX
DE M. tuberculosis and M. leprae marker protein #630.
XX
KW Mycobacterioses; survival; virulence; protective antigen; vaccine;
KW mycobacterial disease; tuberculosis; leprosy.
XX
OS Mycobacterium tuberculosis.
OS Mycobacterium leprae.
XX
PN WO200274903-A2.
XX
PD 26-SEP-2002.
XX
PF 22-FEB-2002; 2002WO-IB001973.
XX
PR 22-FEB-2001; 2001US-0270123P.
XX
PA (INSP) INST PASTEUR.
XX
PI Cole S;
XX
DR WPI; 2002-759885/82.
XX
PT Identifying and selecting genes for survival or virulence of mycobacteria
PT by a comparative genomic analysis of the sequences of Mycobacterium
PT tuberculosis and M. leprae.
XX
PS Claim 17; Page 851-852; 874pp; English.
XX
CC This invention relates to a novel method for identifying essential genes
CC for survival or virulence of mycobacteria species. The method comprises
CC aligning the genomic sequence of a first mycobacterium species on a
CC genomic sequence of a second mycobacterium species and selecting a
CC polynucleotide sequence that is highly conserved in both genomes with no
CC counterparts in other bacterial genomic sequences and that corresponds to
CC an essential gene for the survival or virulence of mycobacterium species.
CC The method of the invention is useful for detecting M. tuberculosis or M.
CC leprae infection. The method reduces the number of potential new targets
CC and protective antigens for new drugs and vaccine compositions to treat
CC and prevent mycobacterial diseases, particularly tuberculosis and
CC leprosy. The present sequence represents a marker protein from
CC Mycobacterium tuberculosis and Mycobacterium leprae identified using the
CC method of the invention
XX
SQ Sequence 637 AA;

Query Match 100.0%; Score 37; DB 5; Length 637;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LAASLLSRV 9
|||||||||
Db 4 LAASLLSRV 12

RESULT 3
AAU65823
ID AAU65823 standard; protein; 175 AA.
XX
AC AAU65823;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #26719.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US012865.
XX
PR 21-APR-2000; 2000US-0199047P.
PR 02-JUN-2000; 2000US-0208841P.
PR 07-JUL-2000; 2000US-0216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59694.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
PS Example 1; SEQ ID NO 27018; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 175 AA;

Query Match 86.5%; Score 32; DB 4; Length 175;
Best Local Similarity 77.8%; Pred. No. 90;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LAASLLSRV 9
|| |||||:
Db 13 LAVSLLSRI 21

RESULT 4
ABM62342
ID ABM62342 standard; protein; 175 AA.
XX
AC ABM62342;
XX
DT 20-OCT-2003 (first entry)
XX
DE Propionibacterium acnes predicted ORF-encoded polypeptide #27018.
XX
KW Acne vulgaris; antiseborrhoeic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine.
XX
OS Propionibacterium acnes.
XX
PN WO2003033515-A1.
XX
PD 24-APR-2003.
XX
PF 11-OCT-2002; 2002WO-US032727.
XX
PR 15-OCT-2001; 2001US-00978825.
XX
PA (CORI-) CORIXA CORP.
XX
PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Vallieve-Douglass J;
XX
DR WPI; 2003-381789/36.
DR N-PSDB; ACF64623.
XX
PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
XX
PS Example 1; SEQ ID NO 27018; 1481pp; English.
XX
CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 175 AA;

Query Match 86.5%; Score 32; DB 6; Length 175;
Best Local Similarity 77.8%; Pred. No. 90;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LAASLLSRV 9
|| |||||:
Db 13 LAVSLLSRI 21

RESULT 5
AAU48527
ID AAU48527 standard; protein; 580 AA.
XX
AC AAU48527;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #9423.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US012865.
XX
PR 21-APR-2000; 2000US-0199047P.
PR 02-JUN-2000; 2000US-0208841P.
PR 07-JUL-2000; 2000US-0216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59543.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
PS Example 1; SEQ ID NO 9722; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as

CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 580 AA;

Query Match 86.5%; Score 32; DB 4; Length 580;
Best Local Similarity 77.8%; Pred. No. 3.4e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LAASLLSRV 9
|| |||||:
Db 67 LAVSLLSRI 75

RESULT 6
ABM45046
ID ABM45046 standard; protein; 580 AA.
XX
AC ABM45046;
XX
DT 20-OCT-2003 (first entry)
XX
DE Propionibacterium acnes predicted ORF-encoded polypeptide #9722.
XX
KW Acne vulgaris; antiseborrhoeic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine.
XX
OS Propionibacterium acnes.
XX
PN WO2003033515-A1.
XX
PD 24-APR-2003.
XX
PF 11-OCT-2002; 2002WO-US032727.
XX
PR 15-OCT-2001; 2001US-00978825.
XX
PA (CORI-) CORIXA CORP.
XX
PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Vallieve-Douglass J;
XX
DR WPI; 2003-381789/36.
DR N-PSDB; ACF64472.
XX
PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
XX
PS Example 1; SEQ ID NO 9722; 1481pp; English.
XX
CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 580 AA;

Query Match 86.5%; Score 32; DB 6; Length 580;
Best Local Similarity 77.8%; Pred. No. 3.4e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LAASLLSRV 9
|| |||||:
Db 67 LAVSLLSRI 75

RESULT 7
AAY82727
ID AAY82727 standard; protein; 382 AA.
XX
AC AAY82727;
XX
DT 19-JUN-2000 (first entry)
XX
DE Thiosulphate sulphurtransferase of wheat.
XX
KW Thiosulphate sulphurtransferase; TS; corn; soybean; wheat; food;
KW foodstuffs; transgenic plants; cyanide; detoxification.
XX
OS Triticum aestivum.
XX
PN WO200006756-A1.
XX
PD 10-FEB-2000.
XX
PF 27-JUL-1999; 99WO-US016920.
XX
PR 28-JUL-1998; 98US-0094453P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Cahoon RE, Falco SC, Rafalski JA;
XX
DR WPI; 2000-205468/18.
DR N-PSDB; AAZ93098.
XX
PT Plant thiosulfate sulfurtransferase genes for generating transgenic
PT plants with altered levels of thiosulfate sulfurtransferase expression
PT useful in feeds for cyanide detoxification.
XX
PS Claim 1; Page 33-34; 39pp; English.
XX
CC Nucleic acids encoding thiosulphate sulphurtransferase genes can be use
CC to generate transgenic plants with altered levels of thiosulphate
CC sulphurtransferase expression. Overexpression of the thiosulphate
CC sulphurtransferase protein in food crops such as soybean, corn and wheat
CC may improve the nutritional value of these crops in foods and feeds. The
CC crops may also be useful for cyanide detoxification. See GENESEQ records
CC AAZ93096-Z93100 and AAY82725-Y82729
XX
SQ Sequence 382 AA;

Query Match 83.8%; Score 31; DB 3; Length 382;
Best Local Similarity 87.5%; Pred. No. 3.4e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LAASLLSR 8
:|||||||

```
Db          1 MAASLLSR 8


RESULT 8
AAY82725
ID   AAY82725 standard; protein; 386 AA.
XX
AC   AAY82725;
XX
DT   19-JUN-2000  (first entry)
XX
DE   Thiosulphate sulphurtransferase of corn.
XX
KW   Thiosulphate sulphurtransferase; TS; corn; soybean; wheat; food;
KW   foodstuffs; transgenic plants; cyanide; detoxification.
XX
OS   Zea mays.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 9
FT                   /note= "Unidentified amino acid"
XX
PN   WO200006756-A1.
XX
PD   10-FEB-2000.
XX
PF   27-JUL-1999;   99WO-US016920.
XX
PR   28-JUL-1998;   98US-0094453P.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Cahoon RE,  Falco SC,  Rafalski JA;
XX
DR   WPI; 2000-205468/18.
DR   N-PSDB; AAZ93096.
XX
PT   Plant thiosulfate sulfurtransferase genes for generating transgenic
PT   plants with altered levels of thiosulfate sulfurtransferase expression
PT   useful in feeds for cyanide detoxification.
XX
PS   Claim 1; Page 30-31; 39pp; English.
XX
CC   Nucleic acids encoding thiosulphate sulphurtransferase genes can be use
CC   to generate transgenic plants with altered levels of thiosulphate
CC   sulphurtransferase expression. Overexpression of the thiosulphate
CC   sulphurtransferase protein in food crops such as soybean, corn and wheat
CC   may improve the nutritional value of these crops in foods and feeds. The
CC   crops may also be useful for cyanide detoxification. See GENESEQ records
CC   AAZ93096-Z93100 and AAY82725-Y82729
XX
SQ   Sequence 386 AA;

  Query Match             83.8%;  Score 31;  DB 3;  Length 386;
  Best Local Similarity   87.5%;  Pred. No. 3.5e+02;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LAASLLSR 8
              :|||||||
Db          1 MAASLLSR 8


RESULT 9
ABP35331
ID   ABP35331 standard; protein; 122 AA.
XX
AC   ABP35331;
XX
DT   09-JUL-2002  (first entry)
XX
DE   Human ORF4304 protein, SEQ ID NO:8608.
XX
KW   Human; ORF; open reading frame; ORFX; drug screening; diagnosis;
KW   disease monitoring; cytokine; cell proliferation; cell differentiation;
KW   immune modulation; haematopoiesis regulation; tissue growth;
KW   angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;
KW   thrombolytic; tumour inhibition; bodily characteristic; fertility;
KW   behaviour; cancer; proliferative disorder; neurological disorder;
KW   cardiovascular disease; immune system disorder; organ transplantation;
KW   tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
KW   hypothyroidism; cholesterol ester storage disease; infection; vulnerary;
KW   vasotropic; antipsoriatic; antidiabetic; cytostatic; nootropic;
KW   neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
KW   cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator;
KW   dermatological; analgesic; virucide; antibacterial; fungicide.
XX
OS   Homo sapiens.
XX
PN   WO200190366-A2.
XX
PD   29-NOV-2001.
XX
PF   24-MAY-2001; 2001WO-US017076.
XX
PR   24-MAY-2000; 2000US-0206690P.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Leach MD,  Shimkets RA;
XX
DR   WPI; 2002-106200/14.
DR   N-PSDB; ABN79357.
XX
PT   Novel human polypeptides and polynucleotides useful for diagnosing,
PT   preventing and treating cardiovascular disease, neurodegenerative,
PT   hyperproliferative disorders and disorders related to organ
PT   transplantation.
XX
PS   Claim 10; Page 2395; 2508pp; English.
XX
CC   Sequences ABP31028-ABP35561 represent 4534 novel human proteins
CC   designated ORF (open reading frame) 1-4534, and sequences ABN75054-
CC   ABN79587 represent cDNAs encoding them. The invention also encompasses
CC   polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
CC   referred to as ORFX) proteins, polynucleotides at least 85% identical to
CC   the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
CC   polynucleotides, the recombinant production of ORFX proteins, antibodies
CC   specific for ORFX proteins, methods of detecting ORFX polynucleotides and
CC   polypeptides, methods of screening for modulators of ORFX expression or
CC   activity, and methods of screening individuals for a predisposition to an
CC   ORFX-associated disorder. The ORFX proteins of the invention have a wide
CC   range of biological activities, such as cytokine, cell proliferation,
CC   cell differentiation, immune modulation, haematopoiesis regulation,
CC   tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
CC   chemokinetic activity, haemostatic activity, thrombolytic activity,
CC   receptor/ligand, antiinflammatory activity, tumour inhibition activity,
CC   and antiinfective activity, and may also be involved in the determination
CC   of bodily characteristics, fertility and behaviour. ORFX proteins,
CC   nucleic acids and antibodies may be used in the treatment of cancers,
CC   other proliferative disorders such as psoriasis and benign tumours,
CC   neurological disorders such as epilepsy and Alzheimer's disease,
CC   cardiovascular diseases, immune system disorders, disorders related to
CC   organ transplantation, disorders of tissue growth and regeneration,
CC   diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
CC   storage disease, and infectious diseases caused by viral, bacterial,
CC   fungal and other pathogens. ORFX nucleic acids may also be used as a
CC   source of primers and probes, in the detection of ORFX genomic sequences
CC   or transcripts, in the identification and cloning of homologous
CC   sequences, in genetic diagnosis, and in forensic biology. The ORFX
CC   nucleic acids may additionally be used to produce transgenic animals
CC   which may be useful for studying the function and/or activity of ORFX
CC   protein, and in drug screening. The ORFX proteins may also be used as
CC   immunogens to generate specific antibodies, which are useful in the
CC   diagnosis, treatment and monitoring of ORFX-associated diseases
XX
SQ   Sequence 122 AA;
```

Query Match 81.1%; Score 30; DB 5; Length 122;
Best Local Similarity 77.8%; Pred. No. 1.6e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LAASLLSRV 9
||||||| :
Db 19 LAASLLSAI 27

RESULT 10
ABM15857
ID ABM15857 standard; protein; 290 AA.
XX
AC ABM15857;
XX
DT 26-SEP-2003 (first entry)
XX
DE Mycobacterium tuberculosis mycobacterial antigen protein SEQ ID NO:79.
XX
KW Mycobacterium tuberculosis; mycobacterial; antigen; infection; vaccine;
KW tuberculostatic; mycobacterial peptide; mycobacterial infection.
XX
OS Mycobacterium tuberculosis.
XX
PN WO2003033530-A2.
XX
PD 24-APR-2003.
XX
PF 14-OCT-2002; 2002WO-GB004647.
XX
PR 12-OCT-2001; 2001GB-00024593.
XX
PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.
XX
PI James B, Bacon J, March P;
XX
DR WPI; 2003-393501/37.
DR N-PSDB; ACF39351.
XX
PT New isolated mycobacterial peptide encoded by a gene that is induced or
PT up-regulated under high oxygen tension, useful for diagnosing, treating
PT or preventing a mycobacterial infection.
XX
PS Claim 1; Page 176-177; 392pp; English.
XX
CC The present invention describes an isolated mycobacterial peptide (I), or
CC its fragment, variant or derivative encoded by a gene whose expression is
CC induced or up-regulated during culture of a mycobacterium under
CC continuous culture conditions of a dissolved oxygen tension of at least
CC 30% air saturation measured at 37 plus degrees Celsius when compared with
CC a dissolved oxygen tension of up to 10% air saturation measured at 37
CC plus degrees Celsius. (I) has tuberculostatic activity and can be used in
CC vaccines. The mycobacterial peptide (I) or its fragment, variant or
CC derivative, inhibitor, antibody, attenuated mycobacterium, attenuated
CC microbial carrier, DNA sequence, DNA plasmid, RNA sequence, or RNA vector
CC from the present invention can be used for manufacturing a medicament for
CC treating or preventing a mycobacterial infection. The peptide or its
CC fragment, variant or derivative, the antibody, or a polynucleotide probe
CC comprising at least 8 nucleotides, where the probe binds to at least a
CC part of the gene, is useful for manufacturing a diagnostic reagent for
CC identifying a mycobacterial infection. The present sequence represents a
CC Mycobacterium tuberculosis mycobacterial antigen, which is used in the
CC exemplification of the present invention
XX
SQ Sequence 290 AA;

Query Match 81.1%; Score 30; DB 6; Length 290;
Best Local Similarity 87.5%; Pred. No. 4.1e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LAASLLSR 8
||||||:|
Db 89 LAASLLAR 96

RESULT 11
AAU35758
ID AAU35758 standard; protein; 422 AA.
XX
AC AAU35758;
XX
DT 14-FEB-2002 (first entry)
XX
DE Helicobacter pylori cellular proliferation protein #71.
XX
KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
KW antibacterial; drug design.
XX
OS Helicobacter pylori.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US009180.
XX
PR 21-MAR-2000; 2000US-0191078P.
PR 23-MAY-2000; 2000US-0206848P.
PR 26-MAY-2000; 2000US-0207727P.
PR 23-OCT-2000; 2000US-0242578P.
PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI; 2001-611495/70.
DR N-PSDB; AAS53617.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
XX
PS Example 3; SEQ ID NO 11351; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes
CC themselves and the encoded proteins. The prokaryotes used are Escherichia
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence represents an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 422 AA;

Query Match 81.1%; Score 30; DB 4; Length 422;
Best Local Similarity 87.5%; Pred. No. 6.2e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LAASLLSR 8
|||:||||
Db 29 LAATLLSR 36

RESULT 12
ABU30835
ID ABU30835 standard; protein; 422 AA.
XX
AC ABU30835;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #16362.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Helicobacter pylori.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR N-PSDB; ACA34705.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 58759; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 422 AA;

Query Match 81.1%; Score 30; DB 6; Length 422;
Best Local Similarity 87.5%; Pred. No. 6.2e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LAASLLSR 8
||| :||||
Db 29 LAATLLSR 36

RESULT 13
AAU62933
ID AAU62933 standard; protein; 509 AA.
XX
AC AAU62933;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #23829.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US012865.
XX
PR 21-APR-2000; 2000US-0199047P.
PR 02-JUN-2000; 2000US-0208841P.
PR 07-JUL-2000; 2000US-0216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59630.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
PS Example 1; SEQ ID NO 24128; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

```
XX
SQ   Sequence 509 AA;

  Query Match             81.1%;  Score 30;  DB 4;  Length 509;
  Best Local Similarity   77.8%;  Pred. No. 7.7e+02;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 LAASLLSRV 9
              ||||||| :
Db        384 LAASLLSAI 392


RESULT 14
ABG05437
ID   ABG05437 standard; protein; 509 AA.
XX
AC   ABG05437;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Novel human diagnostic protein #5428.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   N-PSDB; AAS69624.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 20; SEQ ID NO 35796; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC   amino acid sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 509 AA;

  Query Match             81.1%;  Score 30;  DB 4;  Length 509;
  Best Local Similarity   87.5%;  Pred. No. 7.7e+02;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LAASLLSR 8
              |||:||||
Db        409 LAAALLSR 416


RESULT 15
ABG04461
ID   ABG04461 standard; protein; 509 AA.
XX
AC   ABG04461;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Novel human diagnostic protein #4452.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   N-PSDB; AAS68648.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 20; SEQ ID NO 34820; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC   amino acid sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 509 AA;

  Query Match             81.1%;  Score 30;  DB 4;  Length 509;
```

Best Local Similarity 87.5%; Pred. No. 7.7e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 LAASLLSR 8
              |||:||||
Db        409 LAAALLSR 416
```

Search completed: October 5, 2004, 06:12:55
Job time : 48.9245 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 5, 2004, 06:07:39 ; Search time 26.8302 Seconds
(without alignments)
105.838 Million cell updates/sec

Title: US-10-022-286-5
Perfect score: 42
Sequence: 1 TLLQAAPTL 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	37	88.1	3412	16	Q87PW3	Q87pw3 vibrio para
2	36	85.7	328	16	Q92XV2	Q92xv2 rhizobium m
3	36	85.7	528	16	Q83EA5	Q83ea5 coxiella bu
4	35	83.3	266	16	Q9KBX5	Q9kbx5 bacillus ha
5	35	83.3	998	4	Q7Z5Y3	Q7z5y3 homo sapien
6	35	83.3	1100	11	Q80WC9	Q80wc9 mus musculu
7	34	81.0	255	2	Q9FD88	Q9fd88 bacillus sp
8	34	81.0	1316	16	Q81DQ0	Q81dq0 bacillus ce
9	34	81.0	1925	12	Q9YRB3	Q9yrb3 nudaurelia
10	34	81.0	2385	16	Q81QP7	Q81qp7 bacillus an
11	33	78.6	45	16	Q8CLX2	Q8clx2 yersinia pe
12	33	78.6	101	16	Q8XT67	Q8xt67 ralstonia s
13	33	78.6	353	16	Q8EIL1	Q8eil1 shewanella
14	33	78.6	399	11	Q8BXV8	Q8bxv8 mus musculu
15	33	78.6	459	2	Q50994	Q50994 neisseria g
16	33	78.6	477	16	Q9JZP5	Q9jzp5 neisseria m
17	33	78.6	477	16	Q9JUT1	Q9jut1 neisseria m
18	33	78.6	525	16	Q8Y3W1	Q8y3w1 listeria mo
19	33	78.6	528	5	Q8IRQ6	Q8irq6 drosophila
20	33	78.6	538	10	Q7XVF4	Q7xvf4 oryza sativ
21	33	78.6	595	5	Q8T020	Q8t020 drosophila
22	33	78.6	605	5	Q9W3Y6	Q9w3y6 drosophila
23	33	78.6	627	11	Q8VBY5	Q8vby5 mus musculu
24	33	78.6	2635	12	O40942	O40942 kaposi's sa
25	33	78.6	2635	12	P88955	P88955 kaposi's sa
26	32	76.2	158	16	Q9AJY6	Q9ajy6 streptomyce
27	32	76.2	168	8	Q9B589	Q9b589 lagostomus
28	32	76.2	242	16	Q7WNA8	Q7wna8 bordetella
29	32	76.2	244	16	Q7W0Y4	Q7w0y4 bordetella
30	32	76.2	276	16	Q9KFB5	Q9kfb5 bacillus ha
31	32	76.2	306	16	Q9CJT2	Q9cjt2 pasteurella
32	32	76.2	332	16	Q88EZ5	Q88ez5 pseudomonas
33	32	76.2	522	11	Q8BFW4	Q8bfw4 mus musculu
34	32	76.2	548	2	Q8GAR8	Q8gar8 mycobacteri
35	32	76.2	917	16	Q8G6N4	Q8g6n4 bifidobacte
36	32	76.2	1159	10	Q8LMQ2	Q8lmq2 oryza sativ
37	31	73.8	155	11	Q9DB92	Q9db92 mus musculu
38	31	73.8	176	8	Q9B1E9	Q9b1e9 coryphopter
39	31	73.8	176	8	Q9B1E8	Q9b1e8 coryphopter
40	31	73.8	218	16	Q8ZHZ8	Q8zhz8 yersinia pe
41	31	73.8	266	16	Q7WDQ3	Q7wdq3 bordetella
42	31	73.8	266	16	Q7W2Q5	Q7w2q5 bordetella
43	31	73.8	266	16	Q7W0J7	Q7w0j7 bordetella
44	31	73.8	267	16	Q9A514	Q9a514 caulobacter
45	31	73.8	272	16	Q8NMC3	Q8nmc3 corynebacte

ALIGNMENTS

RESULT 1
Q87PW3
ID Q87PW3 PRELIMINARY; PRT; 3412 AA.
AC Q87PW3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN VP1387.
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RIMD 2210633 / Serotype O3:K6;
RX MEDLINE=22508454; PubMed=12620739;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT distinct from that of V. cholerae.";
RL Lancet 361:743-749(2003).
DR EMBL; AP005077; BAC59650.1; -.
DR GO; GO:0008408; F:3'-5' exonuclease activity; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0004812; F:tRNA ligase activity; IEA.
DR GO; GO:0006418; P:amino acid activation; IEA.
DR GO; GO:0006260; P:DNA replication; IEA.
DR InterPro; IPR006172; DNA_pol_B.
DR InterPro; IPR001412; tRNA-synt_I.
DR PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
DR PROSITE; PS00116; DNA_POLYMERASE_B; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 3412 AA; 363607 MW; F8F0D410E174F547 CRC64;

Query Match 88.1%; Score 37; DB 16; Length 3412;

```
  Best Local Similarity   88.9%;  Pred. No. 2e+02;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TLLQAAPTL 9
              ||||| |||
Db       3005 TLLQALPTL 3013


RESULT 2
Q92XV2
ID   Q92XV2      PRELIMINARY;      PRT;   328 AA.
AC   Q92XV2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Probable ABC transporter, permease protein.
GN   RA1138 OR SMA2085.
OS   Rhizobium meliloti (Sinorhizobium meliloti).
OG   Plasmid pSymA (megaplasmid 1).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX   NCBI_TaxID=382;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1021;
RX   MEDLINE=21396509; PubMed=11481432;
RA   Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
RA   Barloy-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,
RA   Gurjal M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,
RA   Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
RA   Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;
RT   "Nucleotide sequence and predicted functions of the entire
RT   Sinorhizobium meliloti pSymA megaplasmid.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
DR   EMBL; AE007299; AAK65796.1; -.
DR   PIR; B95404; B95404.
DR   GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005215; F:transporter activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR000515; BPD_transp.
DR   Pfam; PF00528; BPD_transp; 1.
DR   PROSITE; PS00402; BPD_TRANSP_INN_MEMBR; 1.
KW   Plasmid; Complete proteome.
SQ   SEQUENCE   328 AA;  35071 MW;  660EEF1B650FFC00 CRC64;

  Query Match             85.7%;  Score 36;  DB 16;  Length 328;
  Best Local Similarity   77.8%;  Pred. No. 28;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TLLQAAPTL 9
              ||||| ||:
Db         14 TLLQAVPTI 22


RESULT 3
Q83EA5
ID   Q83EA5      PRELIMINARY;      PRT;   528 AA.
AC   Q83EA5;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   FrgA protein.
GN   FRGA OR CBU0421.
OS   Coxiella burnetii.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
OC   Coxiellaceae; Coxiella.
OX   NCBI_TaxID=777;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Nine Mile phase I / RSA 493;
RX   MEDLINE=22608657; PubMed=12704232;
RA   Seshadri R., Paulsen I.T., Eisen J.A., Read T.D., Nelson K.E.,
RA   Nelson W.C., Ward N.L., Tettelin H., Davidsen T.M., Beanan M.J.,
RA   DeBoy R.T., Daugherty S.C., Brinkac L.M., Madupu R., Dodson R.J.,
RA   Khouri H.M., Lee K.H., Carty H.A., Scanlan D., Heinzen R.A.,
RA   Thompson H.A., Samuel J.E., Fraser C.M., Heidelberg J.F.;
RT   "Complete genome sequence of the Q-fever pathogen, Coxiella
RT   burnetii.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:5455-5460(2003).
DR   EMBL; AE016961; AAO89973.1; -.
DR   TIGR; CBU0421; -.
DR   GO; GO:0015343; F:siderochrome-iron transporter activity; IEA.
DR   GO; GO:0019290; P:siderochrome biosynthesis; IEA.
DR   InterPro; IPR007310; IucA_IucC.
DR   Pfam; PF04183; IucA_IucC; 1.
KW   Complete proteome.
SQ   SEQUENCE   528 AA;  61066 MW;  83A332AAA9D67D8C CRC64;

  Query Match             85.7%;  Score 36;  DB 16;  Length 528;
  Best Local Similarity   88.9%;  Pred. No. 47;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TLLQAAPTL 9
              |||| ||||
Db        475 TLLQDAPTL 483


RESULT 4
Q9KBX5
ID   Q9KBX5      PRELIMINARY;      PRT;   266 AA.
AC   Q9KBX5;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Oligopeptide ABC transporter (ATP-binding protein).
GN   BH1799.
OS   Bacillus halodurans.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=86665;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C-125 / JCM 9153;
RX   MEDLINE=20512582; PubMed=11058132;
RA   Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA   Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA   Horikoshi K.;
RT   "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT   halodurans and genomic sequence comparison with Bacillus subtilis.";
RL   Nucleic Acids Res. 28:4317-4331(2000).
CC   -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC       (ABC TRANSPORTERS).
DR   EMBL; AP001513; BAB05518.1; -.
DR   PIR; G83874; G83874.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.
DR   GO; GO:0000166; F:nucleotide binding; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR003439; ABC_transporter.
DR   Pfam; PF00005; ABC_tran; 1.
DR   SMART; SM00382; AAA; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR   PROSITE; PS50893; ABC_TRANSPORTER_2; 1.
KW   ATP-binding; Transport; Complete proteome.
SQ   SEQUENCE   266 AA;  29949 MW;  3534AA14F83050AE CRC64;

  Query Match             83.3%;  Score 35;  DB 16;  Length 266;
  Best Local Similarity   87.5%;  Pred. No. 37;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 LLQAAPTL 9
              |:||||||
Db        245 LIQAAPTL 252
```

```
RESULT 5
Q7Z5Y3
ID   Q7Z5Y3      PRELIMINARY;      PRT;   998 AA.
AC   Q7Z5Y3;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Putative non-ribosomal peptide synthetase NRPS998.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Paliga K., Bauer K.;
RT   "Identification and preliminary characterization of a human protein
RT   with similarity to microbial peptide synthetases.";
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF516672; AAP76519.1; -.
SQ   SEQUENCE   998 AA;  111447 MW;  CC8A7409BB1C7A68 CRC64;

  Query Match             83.3%;  Score 35;  DB 4;  Length 998;
  Best Local Similarity   77.8%;  Pred. No. 1.5e+02;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TLLQAAPTL 9
              |:||| |||
Db        184 TVLQATPTL 192


RESULT 6
Q80WC9
ID   Q80WC9      PRELIMINARY;      PRT;  1100 AA.
AC   Q80WC9;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   2-aminoadipic 6-semialdehyde dehydrogenase (EC 1.2.1.31).
GN   AASDH.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kasahara T., Kato T.;
RT   "A new redox-cofactor vitamin for mammals.";
RL   Nature 422:832-832(2003).
DR   EMBL; AB095954; BAC75954.1; -.
DR   GO; GO:0004043; F:L-aminoadipate-semialdehyde dehydrogenase a. . .; IEA.
DR   GO; GO:0016491; F:oxidoreductase activity; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   InterPro; IPR000873; AMP-bind.
DR   InterPro; IPR002372; Bac_PQQ_repeat.
DR   InterPro; IPR006162; Ppantne_S.
DR   InterPro; IPR006163; Pp_bind.
DR   Pfam; PF00501; AMP-binding; 1.
DR   Pfam; PF01011; PQQ; 1.
DR   SMART; SM00564; PQQ; 6.
DR   PROSITE; PS50075; ACP_DOMAIN; 1.
DR   PROSITE; PS00455; AMP_BINDING; 1.
DR   PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.
KW   Oxidoreductase.
SQ   SEQUENCE   1100 AA;  121568 MW;  86B303CFF07B234C CRC64;

  Query Match             83.3%;  Score 35;  DB 11;  Length 1100;
  Best Local Similarity   77.8%;  Pred. No. 1.6e+02;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TLLQAAPTL 9
              |:||| |||
Db        283 TVLQATPTL 291


RESULT 7
Q9FD88
ID   Q9FD88      PRELIMINARY;      PRT;   255 AA.
AC   Q9FD88;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Peptidyl synthetase (Fragment).
OS   Bacillus sp. VK2.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=136090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VK2;
RA   Sokolov S.L., Evdokimova E.G., Esikova T.Z., Temirov Y., Alakhov Y.B.;
RT   "A PCR-based method for identifying of peptidyl synthetase genes from
RT   thermoresistant Bacillus strains.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF289057; AAG03061.1; -.
DR   HSSP; P14687; 1AMU.
DR   GO; GO:0003824; F:catalytic activity; IEA.
DR   GO; GO:0008152; P:metabolism; IEA.
DR   InterPro; IPR000873; AMP-bind.
DR   Pfam; PF00501; AMP-binding; 1.
FT   NON_TER       1      1
FT   NON_TER     255    255
SQ   SEQUENCE   255 AA;  27666 MW;  41998801948E3B3F CRC64;

  Query Match             81.0%;  Score 34;  DB 2;  Length 255;
  Best Local Similarity   66.7%;  Pred. No. 57;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TLLQAAPTL 9
              |::|| |||
Db         88 TIMQATPTL 96


RESULT 8
Q81DQ0
ID   Q81DQ0      PRELIMINARY;      PRT;  1316 AA.
AC   Q81DQ0;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Glycine-AMP ligase (EC 2.7.7.-).
GN   BC2306.
OS   Bacillus cereus (strain ATCC 14579 / DSM 31).
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=226900;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22608415; PubMed=12721630;
RA   Ivanova N., Sorokin A., Anderson I., Galleron N., Candelon B.,
RA   Kapatral V., Bhattacharyya A., Reznik G., Mikhailova N., Lapidus A.,
RA   Chu L., Mazur M., Goltsman E., Larsen N., D'Souza M., Walunas T.,
RA   Grechkin Y., Pusch G., Haselkorn R., Fonstein M., Ehrlich S.D.,
RA   Overbeek R., Kyrpides N.;
RT   "Genome sequence of Bacillus cereus and comparative analysis with
RT   Bacillus anthracis.";
RL   Nature 423:87-91(2003).
DR   EMBL; AE017005; AAP09270.1; -.
DR   GO; GO:0016874; F:ligase activity; IEA.
DR   GO; GO:0016779; F:nucleotidyltransferase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0008152; P:metabolism; IEA.
DR   InterPro; IPR000873; AMP-bind.
DR   InterPro; IPR001242; Condensatn.
DR   InterPro; IPR006163; Pp_bind.
DR   Pfam; PF00501; AMP-binding; 1.
DR   Pfam; PF00668; Condensation; 2.
DR   Pfam; PF00550; pp-binding; 1.
```

```
DR   PROSITE; PS50075; ACP_DOMAIN; 1.
KW   Ligase; Nucleotidyltransferase; Transferase; Complete proteome.
SQ   SEQUENCE   1316 AA;  149211 MW;  08B61F39A1E79DEF CRC64;

  Query Match             81.0%;  Score 34;  DB 16;  Length 1316;
  Best Local Similarity   66.7%;  Pred. No. 3.2e+02;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TLLQAAPTL 9
              |::|| |||
Db        700 TIMQATPTL 708


RESULT 9
Q9YRB3
ID   Q9YRB3      PRELIMINARY;      PRT;  1925 AA.
AC   Q9YRB3;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   RNA-dependent RNA polymerase.
OS   Nudaurelia capensis beta virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Tetraviridae;
OC   Betatetravirus.
OX   NCBI_TaxID=85652;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99263183; PubMed=10329566;
RA   Gordon K.H., Williams M.R., Hendry D.A., Hanzlik T.N.;
RT   "Sequence of the genomic RNA of nudaurelia beta virus (Tetraviridae)
RT   defines a novel virus genome organization.";
RL   Virology 258:42-53(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Gordon K.H.J., Williams M.R., Hendry D.A., Hanzlik T.N.;
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF102884; AAC97509.1; -.
DR   GO; GO:0008174; F:mRNA methyltransferase activity; IEA.
DR   GO; GO:0003724; F:RNA helicase activity; IEA.
DR   GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR   GO; GO:0006396; P:RNA processing; IEA.
DR   GO; GO:0006350; P:transcription; IEA.
DR   GO; GO:0019079; P:viral genome replication; IEA.
DR   InterPro; IPR001788; RNA_dep_RNApol2.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR000606; Viral_helicase1.
DR   InterPro; IPR002588; V_methyltransf.
DR   Pfam; PF00978; RNA_dep_RNApol2; 1.
DR   Pfam; PF01443; Viral_helicase1; 1.
DR   Pfam; PF01660; Vmethyltransf; 1.
KW   RNA-directed RNA polymerase.
SQ   SEQUENCE   1925 AA;  215611 MW;  49A1354A777E9A1A CRC64;

  Query Match             81.0%;  Score 34;  DB 12;  Length 1925;
  Best Local Similarity   77.8%;  Pred. No. 4.7e+02;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TLLQAAPTL 9
              ||||  |||
Db       1799 TLLQGVPTL 1807


RESULT 10
Q81QP7
ID   Q81QP7      PRELIMINARY;      PRT;  2385 AA.
AC   Q81QP7;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Nonribosomal peptide synthetase DhbF.
GN   DHBF OR BA2372.
OS   Bacillus anthracis (strain Ames).
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=198094;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22608414; PubMed=12721629;
RA   Read T.D., Peterson S.N., Tourasse N., Baillie L.W., Paulsen I.T.,
RA   Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,
RA   Holtzapple E.K., Okstad O.A., Helgason E., Rilstone J., Wu M.,
RA   Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.,
RA   DeBoy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H.,
RA   Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,
RA   Benton J.L., Mahamoud Y., Jiang L., Hance I.R., Weidman J.F.,
RA   Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman W.C.,
RA   Hazen A., Cline R., Redmond C., Thwaite J.E., White O., Salzberg S.L.,
RA   Thomason B., Friedlander A.M., Koehler T.M., Hanna P.C., Kolsto A.-B.,
RA   Fraser C.M.;
RT   "The genome sequence of Bacillus anthracis Ames and comparison to
RT   closely related bacteria.";
RL   Nature 423:81-86(2003).
DR   EMBL; AE017031; AAP26240.1; -.
DR   TIGR; BA2372; -.
DR   GO; GO:0003824; F:catalytic activity; IEA.
DR   GO; GO:0016788; F:hydrolase activity, acting on ester bonds; IEA.
DR   GO; GO:0009058; P:biosynthesis; IEA.
DR   InterPro; IPR000873; AMP-bind.
DR   InterPro; IPR001242; Condensatn.
DR   InterPro; IPR006163; Pp_bind.
DR   InterPro; IPR000379; Ser_estrs.
DR   InterPro; IPR001031; Thioesterase.
DR   Pfam; PF00501; AMP-binding; 2.
DR   Pfam; PF00668; Condensation; 2.
DR   Pfam; PF00550; pp-binding; 2.
DR   Pfam; PF00975; Thioesterase; 1.
DR   PRINTS; PR00154; AMPBINDING.
DR   PROSITE; PS50075; ACP_DOMAIN; 2.
DR   PROSITE; PS00455; AMP_BINDING; 1.
KW   Complete proteome.
SQ   SEQUENCE   2385 AA;  267415 MW;  8FE9DA2E3F4CC7FF CRC64;

  Query Match             81.0%;  Score 34;  DB 16;  Length 2385;
  Best Local Similarity   66.7%;  Pred. No. 5.9e+02;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TLLQAAPTL 9
              |::|| |||
Db        700 TIMQATPTL 708


RESULT 11
Q8CLX2
ID   Q8CLX2      PRELIMINARY;      PRT;    45 AA.
AC   Q8CLX2;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical.
GN   Y0017.
OS   Yersinia pestis.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Yersinia.
OX   NCBI_TaxID=632;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KIM5 / Biovar Mediaevalis;
RX   MEDLINE=22137863; PubMed=12142430;
RA   Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA   Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA   Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA   Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA   Perry R.D.;
RT   "Genome sequence of Yersinia pestis KIM.";
RL   J. Bacteriol. 184:4601-4611(2002).
DR   EMBL; AE013603; AAM83614.1; -.
KW   Hypothetical protein.
```

```
SQ   SEQUENCE   45 AA;  4863 MW;  C3AF18AC4A1BB923 CRC64;

  Query Match             78.6%;  Score 33;  DB 16;  Length 45;
  Best Local Similarity   100.0%;  Pred. No. 15;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TLLQAAP 7
              |||||||
Db         17 TLLQAAP 23


RESULT 12
Q8XT67
ID   Q8XT67      PRELIMINARY;      PRT;   101 AA.
AC   Q8XT67;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical protein RSp0248.
GN   RSP0248 OR RS03718.
OS   Ralstonia solanacearum (Pseudomonas solanacearum).
OG   Plasmid megaplasmid.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC   Burkholderiaceae; Ralstonia.
OX   NCBI_TaxID=305;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GMI1000;
RX   MEDLINE=21681879; PubMed=11823852;
RA   Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA   Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA   Chandler M., Choisne N., Claudel-Renard C., Cunnac S., Demange N.,
RA   Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA   Siguier P., Thebault P., Whalen M., Wincker P., Levy M.,
RA   Weissenbach J., Boucher C.A.;
RT   "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL   Nature 415:497-502(2002).
DR   EMBL; AL646077; CAD17399.1; -.
DR   GO; GO:0046821; C:extrachromosomal DNA; IEA.
KW   Plasmid; Hypothetical protein; Complete proteome.
SQ   SEQUENCE   101 AA;  11702 MW;  2B522C4ADBB3CB26 CRC64;

  Query Match             78.6%;  Score 33;  DB 16;  Length 101;
  Best Local Similarity   77.8%;  Pred. No. 35;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TLLQAAPTL 9
              || || |||
Db          5 TLCQAVPTL 13


RESULT 13
Q8EIL1
ID   Q8EIL1      PRELIMINARY;      PRT;   353 AA.
AC   Q8EIL1;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Conserved hypothetical protein.
GN   SO0828.
OS   Shewanella oneidensis.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC   Alteromonadaceae; Shewanella.
OX   NCBI_TaxID=70863;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MR-1;
RX   MEDLINE=22297686; PubMed=12368813;
RA   Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA   Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA   Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA   DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA   Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA   Vamathevan J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,
RA   Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,
RA   Feldblyum T.V., Smith H.O., Venter J.C., Nealson K.H., Fraser C.M.;
RT   "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT   Shewanella oneidensis.";
RL   Nat. Biotechnol. 20:1118-1123(2002).
DR   EMBL; AE015527; AAN53904.1; -.
DR   TIGR; SO0828; -.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0008170; F:N-methyltransferase activity; IEA.
DR   GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf. . .; IEA.
DR   GO; GO:0006306; P:DNA methylation; IEA.
DR   InterPro; IPR007848; MTS.
DR   InterPro; IPR002052; N6_Mtase.
DR   InterPro; IPR000051; SAM_bind.
DR   Pfam; PF05175; MTS; 1.
DR   PROSITE; PS00092; N6_MTASE; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   353 AA;  38476 MW;  5793C391473DAA45 CRC64;

  Query Match             78.6%;  Score 33;  DB 16;  Length 353;
  Best Local Similarity   77.8%;  Pred. No. 1.3e+02;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TLLQAAPTL 9
              |||:| |||
Db        229 TLLKAQPTL 237


RESULT 14
Q8BXV8
ID   Q8BXV8      PRELIMINARY;      PRT;   399 AA.
AC   Q8BXV8;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Ellis van Creveld gene homolog.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK043184; BAC31486.1; -.
DR   PIR; PT0645; PT0645.
DR   PIR; PT0713; PT0713.
SQ   SEQUENCE   399 AA;  44037 MW;  3065516B23DB0C18 CRC64;

  Query Match             78.6%;  Score 33;  DB 11;  Length 399;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LQAAPTL 9
              |||||||
Db         16 LQAAPTL 22


RESULT 15
Q50994
ID   Q50994      PRELIMINARY;      PRT;   459 AA.
AC   Q50994;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Lpd protein.
GN   LPD.
```

```
OS   Neisseria gonorrhoeae.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=485;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MS11;
RA   Porcella S.F., Belland R.J., Judd R.C.;
RT   "The sucAB-lpd operon of Neisseria gonorrhoeae.";
RL   Submitted (MAY-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- COFACTOR: FAD (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE PYRIDINE NUCLEOTIDE-DISULFIDE
CC       OXIDOREDUCTASES CLASS-I.
DR   EMBL; L36381; AAA96487.1; -.
DR   HSSP; P14218; 1LPF.
DR   GO; GO:0005737; C:cytoplasm; IEA.
DR   GO; GO:0004148; F:dihydrolipoamide dehydrogenase activity; IEA.
DR   GO; GO:0015036; F:disulfide oxidoreductase activity; IEA.
DR   GO; GO:0046872; F:metal ion binding; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   GO; GO:0006096; P:glycolysis; IEA.
DR   InterPro; IPR001327; FAD_pyr_redox.
DR   InterPro; IPR000815; Hg_reductase.
DR   InterPro; IPR006258; Lipoamide_dh.
DR   InterPro; IPR000205; NAD_BS.
DR   InterPro; IPR000103; Pyridine_redox_2.
DR   InterPro; IPR001100; Pyr_redox.
DR   InterPro; IPR004099; pyr_redox_dim.
DR   Pfam; PF00070; pyr_redox; 1.
DR   Pfam; PF02852; pyr_redox_dim; 1.
DR   PRINTS; PR00368; FADPNR.
DR   PRINTS; PR00945; HGRDTASE.
DR   PRINTS; PR00411; PNDRDTASEI.
DR   PRINTS; PR00469; PNDRDTASEII.
DR   ProDom; PD000139; FAD_pyr_redox; 1.
DR   TIGRFAMs; TIGR01350; lipoamide_DH; 1.
DR   PROSITE; PS00076; PYRIDINE_REDOX_1; 1.
KW   FAD; Flavoprotein; Oxidoreductase; Redox-active center.
SQ   SEQUENCE   459 AA;  48116 MW;  69685718F4B207D3 CRC64;

  Query Match             78.6%;  Score 33;  DB 2;  Length 459;
  Best Local Similarity   75.0%;  Pred. No. 1.7e+02;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TLLQAAPT 8
              |:|:||||
Db        211 TILEAAPT 218


Search completed: October  5, 2004, 06:17:09
Job time : 32.8302 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 5, 2004, 07:16:11 ; Search time 68 Seconds
(without alignments)
37.396 Million cell updates/sec

Title: US-10-022-286-5
Perfect score: 42
Sequence: 1 TLLQAAPTL 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	42	100.0	9	4	AAB88228	Aab88228 Hsp-65 pe
2	42	100.0	9	5	ABG31249	Abg31249 Immunogen
3	42	100.0	540	1	AAP81351	Aap81351 Sequence
4	42	100.0	540	2	AAR81610	Aar81610 Mycobacte
5	42	100.0	540	2	AAW32100	Aaw32100 Mycobacte
6	42	100.0	540	2	AAW44702	Aaw44702 Mycobacte
7	42	100.0	540	2	AAY23911	Aay23911 Amino aci
8	42	100.0	540	3	AAY93332	Aay93332 Amino aci
9	42	100.0	540	4	AAE11755	Aae11755 Mycobacte
10	42	100.0	540	4	AAG81118	Aag81118 Mycobacte
11	42	100.0	540	4	AAB31606	Aab31606 Amino aci
12	42	100.0	540	5	AAU76511	Aau76511 M. tuberc
13	42	100.0	540	5	AAM50750	Aam50750 Mycobacte
14	42	100.0	540	5	AAU76194	Aau76194 Mycobacte
15	42	100.0	540	6	ABG74588	Abg74588 M. tuberc
16	42	100.0	540	6	ABU34402	Abu34402 Protein e
17	42	100.0	540	6	ABU36456	Abu36456 Protein e
18	42	100.0	540	6	ADA24378	Ada24378 Mycobacte
19	42	100.0	544	2	AAW32099	Aaw32099 Mycobacte
20	42	100.0	560	1	AAP80215	Aap80215 Sequence
21	42	100.0	572	2	AAR04716	Aar04716 Amino aci
22	42	100.0	573	2	AAR64766	Aar64766 M. tuberc
23	42	100.0	583	6	ABR40246	Abr40246 Human rec
24	42	100.0	637	6	ABR40247	Abr40247 Human rec
25	42	100.0	638	3	AAB03790	Aab03790 Heat shoc
26	42	100.0	639	4	AAB31609	Aab31609 Amino aci
27	42	100.0	648	4	AAB31614	Aab31614 Amino aci
28	42	100.0	690	5	ABG70777	Abg70777 BCG Hsp65
29	42	100.0	709	5	ABG70776	Abg70776 BCG Hsp65
30	42	100.0	724	5	ABG70778	Abg70778 BCG Hsp65
31	42	100.0	746	5	ABG70775	Abg70775 His tagge
32	42	100.0	948	4	AAB31611	Aab31611 Amino aci
33	37	88.1	11	4	AAB88263	Aab88263 Hsp-65 pe
34	37	88.1	13	2	AAR97499	Aar97499 Mycobacte
35	37	88.1	15	2	AAR94856	Aar94856 Peptide f
36	37	88.1	16	2	AAW43534	Aaw43534 Mycobacte
37	37	88.1	17	2	AAR94855	Aar94855 Peptide f
38	37	88.1	17	2	AAW43533	Aaw43533 Mycobacte
39	37	88.1	19	2	AAR94854	Aar94854 Peptide f
40	37	88.1	19	2	AAW43532	Aaw43532 Mycobacte
41	37	88.1	540	5	AAU76510	Aau76510 M. leprae
42	37	88.1	540	5	AAU76193	Aau76193 Mycobacte
43	37	88.1	540	6	ABG74587	Abg74587 M. leprae
44	37	88.1	540	6	ADA24377	Ada24377 Mycobacte
45	37	88.1	541	2	AAR67384	Aar67384 M. leprae

ALIGNMENTS

RESULT 1
AAB88228
ID AAB88228 standard; peptide; 9 AA.
XX
AC AAB88228;
XX
DT 17-MAY-2001 (first entry)
XX
DE Hsp-65 peptide epitope #2.
XX
KW Heat shock protein-65; hsp-65; epitope; antibacterial; tuberculostatic;
KW immune response inducer; vaccine; gene therapy; bacterial infection;
KW tuberculosis.
XX
OS Mycobacterium sp.
XX
PN WO200116174-A2.
XX
PD 08-MAR-2001.
XX
PF 30-AUG-2000; 2000WO-IB001326.
XX
PR 30-AUG-1999; 99US-0151396P.
XX
PA (KIES/) KIESSLING R.
PA (CHAR/) CHARO J M.
PA (OTEN/) OTENHOFF T H M.
PA (GELU/) GELUK A.
XX
PI Kiessling R, Charo JM, Otenhoff THM, Geluk A;
XX
DR WPI; 2001-244396/25.
XX
PT Novel polypeptides containing epitopes derived from Mycobacterial heat
PT shock protein 65 useful for treating bacterial and parasitic infections,
PT such as tuberculosis.
XX
PS Claim 1; Page 53; 117pp; English.
XX
CC The present sequence is a peptide epitope derived from Mycobacterial heat
CC shock protein 65. This peptide epitope is useful for inducing a cytotoxic
CC T-cell (CTL) response in vitro for an infectious microbe e.g.
CC Mycobacteria e.g. Mycobacterium tuberculosis; Rickettsia; Chlamydia;
CC Trypanosoma; Helicobacter; Leishmania; Trichomonas e.g. T.vaginalis;
CC Synechoccus e.g. S.vulcanis; Cowdria e.g. C.ruminatium; M.leprae;
CC M.paratuberculosis; Brucella abortus; Leptospira interrogans; Legionella
CC pneumophila; Coxiella burnetii; Staphylococcus aureus; Salmonella typhi;
CC Yersinia enterocolitica; Neisseria meningitidis; N.gonorrhoeae;

CC Haemophilus influenzae and Pseudomonas aeruginosa. This peptide epitope
CC is useful for treating bacterial and parasitic infections such as
CC tuberculosis
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 42; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLLQAAPTL 9
||||||||||
Db 1 TLLQAAPTL 9

RESULT 2
ABG31249
ID ABG31249 standard; peptide; 9 AA.
XX
AC ABG31249;
XX
DT 05-NOV-2002 (first entry)
XX
DE Immunogenic peptide, #3, of M. tuberculosis bovis Hsp65.
XX
KW Immunogen; tuberculosis; TB; tuberculostatic; HLA-A2;
KW human leukocyte antigen A2; major histocompatibility complex; MHC;
KW cytotoxic T lymphocyte; CTL; cytokine; immunotherapy; epitope; Hsp65;
KW therapeutic; MHC tetramer; lymphocyte; bovis heat shock protein 65.
XX
OS Mycobacterium tuberculosis.
XX
PN WO200248175-A2.
XX
PD 20-JUN-2002.
XX
PF 12-DEC-2001; 2001WO-US048742.
XX
PR 13-DEC-2000; 2000US-0255292P.
PR 30-JAN-2001; 2001US-0264978P.
XX
PA (ARGO-) ARGONEX PHARM.
PA (UYVI-) UNIV VIRGINIA PATENT FOUND.
XX
PI Flyer D, Ross MM, Hunt DF, White FM;
XX
DR WPI; 2002-599500/64.
XX
PT Novel immunogen comprising a peptide segment derived from Mycobacterium
PT tuberculosis, useful for inducing a cytotoxic T lymphocyte response in
PT vivo to tuberculosis infected cells.
XX
PS Claim 2; Page 45; 52pp; English.
XX
CC The invention discloses an immunogen, and polynucleotide encoding it,
CC comprising a peptide segment derived from Mycobacterium tuberculosis,
CC where the immunogen in not heat shock protein (hsp) 65 protein. The
CC peptide, and antibody specific for the peptide, can be used to treat a
CC subject with tuberculosis (TB) characterised by tuberculosis infected
CC cells expressing human leukocyte antigen A2 (HLA-A2) or any class I major
CC histocompatibility complex (MHC) molecule. The treatment involves
CC administering cytotoxic T lymphocytes (CTLs), induced in vitro using the
CC peptide, to destroy infected cells through direct lysis or to effect the
CC destruction of the infected cells indirectly through the elaboration of
CC cytokines. The ability of the immunogens to generate CTLs in vitro can
CC serve as a diagnostic for tuberculosis and they can be useful for
CC stimulating production of antibodies for use in passive immunotherapy.
CC The peptides are useful for screening a sample for the presence of CTLs
CC that specifically recognise the corresponding epitopes, as a diagnostic
CC tool to evaluate the therapeutic efficacy of the immunotherapeutic
CC treatments and to prepare class I MHC tetramers which can be used in
CC conjunction with flow cytometry to quantitate the frequency of peptide-
CC specific CTLs that are present in a sample of lymphocytes from an
CC individual. The sequence presented is the immunogenic peptide, #3, of the
CC M. tuberculosis bovis Hsp65
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 42; DB 5; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLLQAAPTL 9
||||||||||
Db 1 TLLQAAPTL 9

RESULT 3
AAP81351
ID AAP81351 standard; protein; 540 AA.
XX
AC AAP81351;
XX
DT 25-MAR-2003 (revised)
DT 29-DEC-1990 (first entry)
XX
DE Sequence of Mycobacterium tuberculosis 540 AA residue protein.
XX
KW Diagnosis; assay; M.bovis; vaccine.
XX
OS Mycobacterium tuberculosis.
XX
PN WO8806591-A.
XX
PD 07-SEP-1988.
XX
PF 25-FEB-1988; 88WO-US000598.
XX
PR 26-FEB-1987; 87US-00019529.
PR 24-FEB-1988; 88US-00159667.
XX
PA (SCRI) SCRIPPS CLINIC & RE.
XX
PI Shinnick T, Houghten R;
XX
DR WPI; 1988-271136/38.
DR N-PSDB; AAN81768.
XX
PT Recombinant mycobacterial peptide(s) - used in assays for diagnosis of
PT infection, for producing vaccines and for producing antibodies.
XX
PS Disclosure; Fig 2a-2d; 117pp; English.
XX
CC An isolated DNA molecule that consists essentially of the nucleotide
CC sequence that corresponds to the sequence represented by position 3950 to
CC about 2390 and from position 3948 through position 2398 of AAN81768 is
CC claimed. Also claimed is a peptide sequence that consists of a 5-40 AA
CC residue sequence that corresponds to a sequence of the 540 AA residue
CC protein (AAP81351) or the 517 AA residue protein (AAP81868) coded for by
CC the DNA sequence. The proteins can be used for determining previous
CC immunological exposure of a mammal to M.tuberculosis or M.bovis and for
CC producing a vaccine. (Updated on 25-MAR-2003 to correct PR field.)
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 540 AA;

Query Match 100.0%; Score 42; DB 1; Length 540;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLLQAAPTL 9
||||||||||
Db 416 TLLQAAPTL 424

RESULT 4

AAR81610
ID AAR81610 standard; protein; 540 AA.
XX
AC AAR81610;
XX
DT 09-MAY-1996 (first entry)
XX
DE Mycobacterium tuberculosis heat shock protein hsp65.
XX
KW Heat shock protein; mycobacterium tuberculosis; inflammatory disease;
KW autoimmune disease; diabetes; arthritis; atherosclerosis; antibody;
KW multiple sclerosis; myasthemia gravis; transplant rejection; diagnosis;
KW therapy.
XX
OS Mycobacterium tuberculosis.
XX
PN WO9525744-A1.
XX
PD 28-SEP-1995.
XX
PF 21-MAR-1995; 95WO-NL000108.
XX
PR 21-MAR-1994; 94EP-00200721.
PR 22-MAR-1994; 94EP-00200738.
PR 10-OCT-1994; 94EP-00202927.
XX
PA (UYUT-) RIJKSUNIV UTRECHT.
XX
PI Anderton SM, Van Der Zee R, Van Eden W;
XX
DR WPI; 1995-344587/44.
XX
PT Microbial stress protein fragments containing epitope(s) homologous to
PT related mammalian epitope(s) - used to treat and prevent inflammation
PT e.g. auto:immune diseases, also nucleic acids, vectors and recombinant
PT cells.
XX
PS Claim 4; Fig 13; 65pp; English.
XX
CC This sequence represents the heat shock protein hsp65 of Mycobacterium
CC tuberculosis. The peptide fragments of this sequence represented by
CC AAR81611-R81623 were used to immunise rat T-cells. These peptide
CC fragments represent regions of stress proteins that are highly conserved
CC between microorganisms and animals. The immunisation was carried out in
CC order to protect the rat from inflammatory diseases. The inflammatory
CC diseases that the peptides can be used to treat include autoimmune
CC diseases such as diabetes, arthritis, atherosclerosis, multiple sclerosis
CC and myasthemia gravis. They also prevent transplant rejection. Antibodies
CC raised against the peptide sequences can be used in diagnosis, e.g. to
CC measure expression of the epitopes at sites of inflammation, or to
CC measure T-cell proliferation or cytokine production. The antibodies can
CC also be used for passive immunisation. The peptide fragments
CC corresponding to similar regions of mammalian stress proteins do not
CC elicit a protective response
XX
SQ Sequence 540 AA;

Query Match 100.0%; Score 42; DB 2; Length 540;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLLQAAPTL 9
|||||||||
Db 416 TLLQAAPTL 424

RESULT 5
AAW32100
ID AAW32100 standard; protein; 540 AA.
XX
AC AAW32100;
XX
DT 27-AUG-2003 (revised)
DT 08-APR-1998 (first entry)
XX
DE Mycobacteria sp. heat shock protein 65 (hsp65) #2.
XX
KW Heat shock protein; hsp58; human; autoimmune disease; hsp65;
KW rheumatoid arthritis; antigen; infectious disease; prophylactic;
KW pristane induced arthritis; PIA; vaccine.
XX
OS Mycobacteriaceae.
XX
PN WO9711966-A1.
XX
PD 03-APR-1997.
XX
PF 26-SEP-1996; 96WO-GB002382.
XX
PR 27-SEP-1995; 95GB-00019737.
XX
PA (PEPT-) PEPTIDE THERAPEUTIC LTD.
XX
PI Thompson JS, Elson CJ;
XX
DR WPI; 1997-212851/19.
XX
PT Polypeptide(s) derived from microbial heat shock protein - useful for
PT treatment of autoimmune disease esp. arthritis.
XX
PS Disclosure; Fig 4; 91pp; English.
XX
CC This is the heat shock protein hsp65 which can be used in a novel method
CC to treat autoimmune disease e.g. rheumatoid arthritis. This sequence is
CC known to be an immunodominant antigen in a number of infectious diseases
CC and is linked to pristane induced arthritis (PIA) in vitro. HSP's from
CC microbial sources may act as self antigens and thus have limited use
CC whereas the human hsp65 homologue, hsp58 or fragments of the hsp58
CC protein may be useful in the development of vaccines for prophylaxis or
CC treatment of an autoimmune disease such as rheumatoid arthritis. Note:
CC this hsp65 sequence is represented in Figure 4 but differs slightly from
CC the hsp65 sequence given in the sequence specification (see AAW32099).
CC (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ Sequence 540 AA;

Query Match 100.0%; Score 42; DB 2; Length 540;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLLQAAPTL 9
|||||||||
Db 416 TLLQAAPTL 424

RESULT 6
AAW44702
ID AAW44702 standard; protein; 540 AA.
XX
AC AAW44702;
XX
DT 22-JUN-1998 (first entry)
XX
DE Mycobacterium tuberculosis 65 kDa heat shock protein (Hsp65).
XX
KW Heat shock protein; Mt Hsp65; autoimmune disease; immunotherapy;
KW gene therapy; rheumatoid arthritis; multiple sclerosis.
XX
OS Mycobacterium tuberculosis.
XX
PN WO9746253-A2.
XX
PD 11-DEC-1997.
XX
PF 03-JUN-1997; 97WO-US009427.
XX

PR 03-JUN-1996; 96US-0019100P.
PR 03-JUN-1997; 97US-00019100.
XX
PA (AURA-) AURAGEN INC.
XX
PI Haynes JR, Prayaga SK, Ramshaw IA;
XX
DR WPI; 1998-041892/04.
DR N-PSDB; AAV05708.
XX
PT Treatment of auto-immune diseases - by administering auto-antigen-coated
PT particles or auto-antigen-encoding nucleic acid construct.
XX
PS Example 2; Page 55-58; 72pp; English.
XX
CC This protein comprises the 65 kDa heat shock protein, Mt Hsp65, of
CC Mycobacterium tuberculosis. This protein cross-reacts with a component of
CC articular cartilage, human Hsp60, that is up-regulated in the joints of
CC arthritic patients. The Mt Hsp65 gene sequence (see AAV05708) is also
CC provided. A claimed method for treating or preventing an autoimmune
CC disease in a mammal comprises: (a) providing a particle coated with an
CC antigen against which an immune response is mounted in the autoimmune
CC disease; (b) delivering the particle into the recipient cell of the
CC mammal; and (c) repeating step (b) until either a reduction in a
CC cytotoxic immune response or a desensitizing immune response is induced
CC in the mammal. Alternatively, step (a) comprises providing a nucleic acid
CC construct comprising a coding sequence for the antigen, operably linked
CC to control elements such that the coding sequence can be transcribed and
CC translated in a recipient cell, and delivering the construct to the
CC recipient cell using a gene gun. The antigen of step (a) is selected from
CC collagen, Mt Hsp65, myelin basic protein, myelin oligodendrocyte
CC glycoprotein, proteolipid protein, and epitopes thereof. These antigens
CC mitigate cytotoxic responses and elicit antigen desensitisation. The
CC method is used especially for treating rheumatoid arthritis or multiple
CC sclerosis. It represents a novel use for the previously known Mt Hsp65
XX
SQ Sequence 540 AA;

Query Match 100.0%; Score 42; DB 2; Length 540;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLLQAAPTL 9
|||||||||
Db 416 TLLQAAPTL 424

RESULT 7
AAY23911
ID AAY23911 standard; protein; 540 AA.
XX
AC AAY23911;
XX
DT 22-SEP-1999 (first entry)
XX
DE Amino acid sequence of a heat shock protein.
XX
KW Heat shock protein; Hsp; immune response; immunological carrier;
KW cancer control; tumour; sarcoma; cancer; gene therapy.
XX
OS Mycobacterium bovis.
XX
PN WO9935270-A1.
XX
PD 15-JUL-1999.
XX
PF 29-DEC-1998; 98WO-CA001203.
XX
PR 31-DEC-1997; 97US-00001737.
XX
PA (STRE-) STRESSGEN BIOTECHNOLOGIES CORP.
XX
PI Mizzen L, Wisniewski J;
XX
DR WPI; 1999-430397/36.
XX
PT New nucleic acid encoding heat shock protein-60 from Streptococcus,
PT useful in vaccines, as carriers for other immunogens, as anticancer
PT agents and for diagnosis.
XX
PS Disclosure; Fig 10A-E; 176pp; English.
XX
CC AAY23905-30 represent heat shock proteins (Hsps). The specification
CC describes Streptococcal Hsps, designated Hsp60. These proteins, their
CC fragments, variants and fusion proteins, are used to elicit or enhance an
CC immune response against Streptococcus, and to elicit a similar response
CC to a target antigen fused to the protein. Unlike other immunological
CC carriers, Hsp60 proteins are not immunosuppressive so provide an
CC increased response to any conjugated or fused antigen. Also, where used
CC for cancer control, they lack the side effects associated with
CC endotoxins. They can also be used to detect specific antibodies and in
CC treatment or prevention of tumours (e.g. sarcoma or cancers of breast,
CC ovary, prostate, lung, pancreas or liver). The Hsp60 polynucleotide is
CC used for recombinant production of the protein, as a source of primers
CC and probes for detecting streptococci in standard
CC hybridization/amplification assays, and therapeutically in gene therapy
CC vectors
XX
SQ Sequence 540 AA;

Query Match 100.0%; Score 42; DB 2; Length 540;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLLQAAPTL 9
|||||||||
Db 416 TLLQAAPTL 424

RESULT 8
AAY93332
ID AAY93332 standard; peptide; 540 AA.
XX
AC AAY93332;
XX
DT 04-SEP-2000 (first entry)
XX
DE Amino acid sequence of a heat shock protein 60.
XX
KW Epitope; heat shock protein 60; Hsp60; vaccine; autoimmune disease;
KW inflammatory disorder; arthritis.
XX
OS Mycobacterium tuberculosis.
XX
PN WO200027870-A1.
XX
PD 18-MAY-2000.
XX
PF 04-NOV-1999; 99WO-IL000595.
XX
PR 05-NOV-1998; 98US-0107213P.
XX
PA (HADA-) HADASIT MEDICAL RES SERVICES & DEV.
XX
PI Naparstek Y, Ulmansky R, Kashi Y;
XX
DR WPI; 2000-376486/32.
XX
PT Peptide having a defined sequence is used in vaccines for conferring
PT immunity against autoimmune disease or inflammatory disorders, especially
PT arthritis.
XX
PS Disclosure; Fig 1; 58pp; English.
XX
CC The present sequence represents a heat shock protein 60 (Hsp60) of
CC Mycobacterium tuberculosis. The specification describes epitopes of Hsp60

CC proteins, which are useful in vaccines for conferring immunity against
CC autoimmune disease or inflammatory disorders, especially arthritis. The
CC peptide may also be used to raise antibodies, which are then used for
CC passive immunisation
XX
SQ Sequence 540 AA;

Query Match 100.0%; Score 42; DB 3; Length 540;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLLQAAPTL 9
|||||||||
Db 416 TLLQAAPTL 424

RESULT 9
AAE11755
ID AAE11755 standard; protein; 540 AA.
XX
AC AAE11755;
XX
DT 18-DEC-2001 (first entry)
XX
DE Mycobacterium tuberculosis heat shock protein 65 (HSP65).
XX
KW Heat shock protein 65; HSP65; antiarteriosclerotic; antiinflammatory;
KW antiallergic; immunomodulator; dermatological; immunosuppressive;
KW vasoactive; immunostimulant; therapy; vascular disorder; immune response;
KW atherosclerosis; allergic angiitis; Behcet's syndrome; granulomatosis;
KW Churg-Strauss disease; Cogan's syndrome; graft-versus-host disease; GvHD;
KW Henoch-Schonlein purpura; leucocytoclastic vasculitis; Kawasaki disease;
KW polyarteritis nodosa; PAN; Takayasu's arteritis; temporal arteritis;
KW polyangiitis overlap syndrome; Buerger's disease; transplant rejection;
KW thromboangiitis obliterans; Wegener's granulomatosis;
KW microscopic polyangiitis.
XX
OS Mycobacterium tuberculosis.
XX
PN WO200168124-A2.
XX
PD 20-SEP-2001.
XX
PF 15-MAR-2001; 2001WO-US008351.
XX
PR 15-MAR-2000; 2000US-0189855P.
XX
PA (BGHM) BRIGHAM & WOMENS HOSPITAL INC.
XX
PI Weiner HL, Maron R, Libby P;
XX
DR WPI; 2001-611383/70.
XX
PT Treating a vascular disorder, involves administering a composition
PT comprising heat shock protein, its fragment or analog, by mucosal
PT surface, pulmonary tract, oral or enteral route, or by inhalation.
XX
PS Disclosure; Page 11; 49pp; English.
XX
CC The patent discloses methods for treating vascular disorders in mammals.
CC The method involves administering a composition comprising at least one
CC agent selected from heat shock protein (HSP), its fragment or analogue,
CC through mucosal surface, pulmonary tract, oral or enteral route or by
CC inhalation. Compositions comprising HSP are useful for treating and
CC suppressing a vascular disorder, including cell-mediated immune response,
CC an antibody-mediated immune response, cell-mediated inflammatory
CC disorder, atherosclerosis, allergic angiitis, Behcet's syndrome,
CC granulomatosis (Churg-Strauss disease), Cogan's syndrome, graft-versus-
CC host disease (GvHD), Henoch-Schonlein purpura, Kawasaki disease,
CC leucocytoclastic vasculitis, polyarteritis nodosa (PAN), microscopic
CC polyangiitis, polyangiitis overlap syndrome, Takayasu's arteritis,
CC temporal arteritis, transplant rejection, Wegener's granulomatosis and
CC thromboangiitis obliterans (Buerger's disease). They are useful for
CC reducing the level of proinflammatory Th1 cytokines and also for
CC increasing the level of antiinflammatory Th2 cytokines. The present
CC sequence is heat shock protein 65 (HSP65) from Mycobacterium tuberculosis
XX
SQ Sequence 540 AA;

Query Match 100.0%; Score 42; DB 4; Length 540;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLLQAAPTL 9
|||||||||
Db 416 TLLQAAPTL 424

RESULT 10
AAG81118
ID AAG81118 standard; protein; 540 AA.
XX
AC AAG81118;
XX
DT 04-SEP-2001 (first entry)
XX
DE Mycobacterium tuberculosis potential drug target protein SEQ ID 169.
XX
KW Drug target; growth; organism viability; characterisation.
XX
OS Mycobacterium tuberculosis.
XX
PN WO200135317-A1.
XX
PD 17-MAY-2001.
XX
PF 13-NOV-2000; 2000WO-US031152.
XX
PR 12-NOV-1999; 99US-0165086P.
PR 12-NOV-1999; 99US-0165124P.
PR 01-FEB-2000; 2000US-0179531P.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Eisenberg D, Rotstein SH, Marcotte EM;
XX
DR WPI; 2001-329193/34.
DR N-PSDB; AAH51969.
XX
PT Identifying nucleotide or polypeptide sequence for use as drug target,
PT involves providing algorithm that analyzes a functional relationship
PT between nucleotide or polypeptide sequences, and comparing the sequences.
XX
PS Disclosure; Page 160; 207pp; English.
XX
CC This invention relates to a method for identifying a nucleotide or
CC polypeptide sequence that may be a drug target, or essential for growth
CC or viability of an organism. Polynucleotide sequences AAH51947 - AAH52092
CC represent DNA encoding proteins AAG81096 - AAG81241, Mycobacterium
CC tuberculosis proteins which are potential drug targets. The DNA and
CC protein sequences are used to illustrate the method of the invention. The
CC method involves providing an unknown nucleotide or polypeptide sequences,
CC and comparing it to a number of sequences along with at least one
CC algorithm capable of analysing a functional relationship between
CC nucleotide and polypeptide sequences. The method is useful for
CC characterising the function of nucleic acids and polypeptides that may be
CC useful as a target for a drug or essential for the growth or viability of
CC an organism
XX
SQ Sequence 540 AA;

Query Match 100.0%; Score 42; DB 4; Length 540;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLLQAAPTL 9

```
                  |||||||||
Db        416 TLLQAAPTL 424


RESULT 11
AAB31606
ID   AAB31606 standard; protein; 540 AA.
XX
AC   AAB31606;
XX
DT   30-APR-2001  (first entry)
XX
DE   Amino acid sequence of M. bovis BCG heat shock protein 65 (Hsp65).
XX
KW   Heat shock protein; Hsp; Th1 response; Th1 cell; CD4+ T lymphocyte cell;
KW   lymphocyte; Hsp65; Hsp40; Hsp10; Hsp60; Hsp71; microbial pathogen.
XX
OS   Mycobacterium bovis.
XX
PN   WO200104344-A2.
XX
PD   18-JAN-2001.
XX
PF   10-JUL-2000; 2000WO-US018828.
XX
PR   08-JUL-1999;   99US-0143757P.
XX
PA   (STRE-) STRESSGEN BIOTECHNOLOGIES CORP.
XX
PI   Siegel M,  Chu NR,  Mizzen LA;
XX
DR   WPI; 2001-138361/14.
DR   N-PSDB; AAF25002.
XX
PT   Screening for compounds that stimulate Th1-like responses in CD4+ T
PT   lymphocyte cells.
XX
PS   Example 2; Fig 1A-B; 88pp; English.
XX
CC   The present sequence represents the Mycobacterium bovis BCG heat shock
CC   protein (Hsp) 65. Hsp65 is used in the course of the invention. The
CC   specification describes a method of determining whether a compound
CC   stimulates a Th1-like response. Th1 cells are a subset of CD4+ T
CC   lymphocyte cells. The method comprises contacting naive lymphocytes in
CC   vitro with a fusion protein comprising at least a fragment of Hsp, and
CC   then detecting the TH1-like response exhibited by the cell sample. The
CC   proteins which may be used in the method of the invention are Hsp65,
CC   Hsp40, Hsp10, Hsp60, and Hsp71. The method may be used to identify
CC   compounds that stimulate Th1-like responses in response to microbial
CC   pathogens
XX
SQ   Sequence 540 AA;

  Query Match             100.0%;  Score 42;  DB 4;  Length 540;
  Best Local Similarity   100.0%;  Pred. No. 4.8;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TLLQAAPTL 9
              |||||||||
Db        416 TLLQAAPTL 424


RESULT 12
AAU76511
ID   AAU76511 standard; protein; 540 AA.
XX
AC   AAU76511;
XX
DT   05-JUN-2002  (first entry)
XX
DE   M. tuberculosis 65kDA stress protein.
XX
KW   Virucide; anti-HIV; cytostatic; stress protein; immune response;
KW   human immunodeficiency virus p24 protein; HIV; cancer;
KW   Neisseria meningitidis group B; Streptococcus pneumoniae type 14.
XX
OS   Mycobacterium tuberculosis.
XX
PN   US6338952-B1.
XX
PD   15-JAN-2002.
XX
PF   03-NOV-1994;   94US-00336251.
XX
PR   15-JUN-1988;   88US-00207298.
PR   15-JUN-1989;   89US-00366581.
PR   15-JUN-1989;   89WO-US002619.
PR   09-DEC-1991;   91US-00804632.
PR   04-JUN-1993;   93US-00073381.
PR   06-JUN-1994;   94WO-US006362.
XX
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX
PI   Young RA;
XX
DR   WPI; 2002-215020/27.
XX
PT   New isolated fusion protein, comprises stress protein fused to
PT   heterologous protein e.g. cancer antigen, useful for inducing or
PT   enhancing immune response against heterologous protein when administered
PT   to individual.
XX
PS   Example 1; Fig 4; 29pp; English.
XX
CC   The invention relates to an isolated fusion protein (I) comprising a
CC   stress protein joined via a peptide bond to an heterologous protein or
CC   peptide, where (I) when administered to an individual induces or enhances
CC   immune response against the heterologous protein or peptide. (I) is
CC   useful for inducing or enhancing immune response in an individual,
CC   against viral (human immunodeficiency virus p24 protein or peptide), or a
CC   cancer antigen. (I) is useful for inducing an immune response against
CC   oligosaccharides of bacteria such as Neisseria meningitidis group B and
CC   Streptococcus pneumoniae type 14. The present sequence represents the
CC   amino acid sequence of Mycobacterium tuberculosis 65kDa stress protein of
CC   the invention
XX
SQ   Sequence 540 AA;

  Query Match             100.0%;  Score 42;  DB 5;  Length 540;
  Best Local Similarity   100.0%;  Pred. No. 4.8;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TLLQAAPTL 9
              |||||||||
Db        416 TLLQAAPTL 424


RESULT 13
AAM50750
ID   AAM50750 standard; protein; 540 AA.
XX
AC   AAM50750;
XX
DT   18-APR-2002  (first entry)
XX
DE   Mycobacterium tuberculosis immunodominant Mtb protein GroEL2.
XX
KW   Mtb; GroEL2; immunogen; mycobacteria; immunisation; vaccine.
XX
OS   Mycobacterium tuberculosis.
XX
PN   WO200204018-A2.
XX
PD   17-JAN-2002.
XX
PF   10-JUL-2001; 2001WO-US021717.
```

XX
PR 10-JUL-2000; 2000US-0217646P.
XX
PA (COLS) UNIV COLORADO STATE RES FOUND.
XX
PI Orme IM, Belisle JT;
XX
DR WPI; 2002-164602/21.
XX
PT Vaccine for boosting immunity to mycobacteria when administered in mid-
PT life in a subject who has been vaccinated in childhood with Bacillus
PT Calmette-Guerrin, has purified proteins from mycobacterium tuberculosis.
XX
PS Claim 8; Page 19; 61pp; English.
XX
CC The present sequence is that of the Mycobacterium tuberculosis (Mtb)
CC strain H37Rv cytosolic Rv0440 gene product, designated GroEL2. This is
CC one of 31 immunodominant secreted or cytosolic Mtb proteins of strain
CC H37Rv (see AAM50729-59) discovered through the use of 2-dimensional
CC liquid phase electrophoresis coupled with an in vitro interferon-gamma
CC assay and liquid chromatography-mass spectrometry. The immunogens
CC stimulate a strong interferon-gamma response from T cells of M.
CC tuberculosis infected mice. The invention provides vaccine compositions
CC for boosting immunity to mycobacteria when administered in mid-life to a
CC subject who has been vaccinated neonatally or in early childhood with BCG
CC and in whom protective immunity has waned. The vaccine compositions
CC comprise 1 or more of the 31 purified immunogenic proteins. A preferred
CC protein is Ag85A (see AAM50759), the secreted product of the Rv3084v gene
XX
SQ Sequence 540 AA;

Query Match 100.0%; Score 42; DB 5; Length 540;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLLQAAPTL 9
|||||||||
Db 416 TLLQAAPTL 424

RESULT 14
AAU76194
ID AAU76194 standard; protein; 540 AA.
XX
AC AAU76194;
XX
DT 08-MAY-2002 (first entry)
XX
DE Mycobacterium tuberculosis 65 kDa antigen stress protein (TB65K).
XX
KW TB65K virucide; anti-HIV; human immunodeficiency virus; antibacterial;
KW antiparasitic; immunosuppressive; antiarthritic; antirheumatic;
KW stress protein; viral antigen; gag; pol; p24 protein; cancer; autoimmune;
KW rheumatoid arthritis.
XX
OS Mycobacterium tuberculosis.
XX
PN US6335183-B1.
XX
PD 01-JAN-2002.
XX
PF 05-JUN-1995; 95US-00461722.
XX
PR 15-JUN-1988; 88US-00207298.
PR 15-JUN-1989; 89US-00366581.
PR 15-JUN-1989; 89WO-US002619.
PR 09-DEC-1991; 91US-00804632.
PR 04-JUN-1993; 93US-00073381.
PR 06-JUN-1994; 94WO-US006362.
PR 03-NOV-1994; 94US-00336251.
XX
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
XX
PI Young RA, Young D;
XX
DR WPI; 2002-163203/21.
XX
PT Inducing or enhancing immune response in a patient by administering to
PT the patient, a composition comprising an isolated fusion protein which
PT comprises a stress protein or its part, fused to heterologous
PT polypeptide.
XX
PS Example 1; Fig 4; 29pp; English.
XX
CC This invention relates to a novel method for inducing or enhancing immune
CC responses in a patient. The method comprises administering to the patient
CC a pharmaceutical composition comprising an isolated fusion protein having
CC a stress protein or a portion of stress protein fused to a heterologous
CC protein or peptide. The fusion protein, when administered to the patient,
CC induces or enhances immune response against the heterologous protein or
CC peptide. The method of the invention may be used for inducing or
CC enhancing immune response in a patient against a heterologous protein or
CC peptide which is administered as a part of the fusion protein e.g. viral
CC antigen such as an human immunodeficiency virus (HIV) protein or peptide
CC e.g. gag or pol protein or peptide, preferably p24 protein or peptide, or
CC a cancer antigen. The method is also useful for inducing or enhancing an
CC individuals immune response to other pathogens such as bacteria,
CC parasites, or other organisms or agents such as toxins or toxoids. It is
CC also useful for enhancing or inducing an upregulation of an individual's
CC immune status (such as in an immune compromised individual or HIV-
CC infected individual), and to decrease an individual's autoimmune response
CC such as that which occurs in rheumatoid arthritis. The administration of
CC the stress protein also provides protection against subsequent infection
CC by a pathogen. The present sequence represents the mycobacterium
CC tuberculosis 65kDa antigen (TB65K) stress protein used in the method of
CC the invention
XX
SQ Sequence 540 AA;

Query Match 100.0%; Score 42; DB 5; Length 540;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLLQAAPTL 9
|||||||||
Db 416 TLLQAAPTL 424

RESULT 15
ABG74588
ID ABG74588 standard; protein; 540 AA.
XX
AC ABG74588;
XX
DT 13-MAY-2003 (first entry)
XX
DE M. tuberculosis 65kDa stress protein.
XX
KW Stress protein; immune response; heat shock protein; haemagglutinin;
KW hsp90; hsp70; hsp60; groES; DnaJ; DnaK; GroEL; Forssman antigen; HIV;
KW cancer cell-associated antigen; GD3; GM2; Gb3; Sialosyl-Lea; Tn antigen;
KW melanoma antigen; carcino-embryonic antigen; CEA; alpha-fetoprotein;
KW prostate specific antigen; viral antigen; human immunodeficiency virus;
KW gag; pol; parasitic antigen.
XX
OS Mycobacterium tuberculosis.
XX
PN US6482614-B1.
XX
PD 19-NOV-2002.
XX
PF 21-DEC-1999; 99US-00468041.
XX
PR 15-JUN-1988; 88US-00207298.
PR 15-JUN-1989; 89US-00366581.
PR 15-JUN-1989; 89WO-US002619.

PR 09-DEC-1991; 91US-00804632.
PR 04-JUN-1993; 93US-00073381.
PR 06-JUN-1994; 94WO-US006362.
PR 03-NOV-1994; 94US-00336251.
PR 05-JUN-1995; 95US-00461720.
XX
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
XX
PI Young RA;
XX
DR WPI; 2003-298137/29.
XX
PT Producing antibodies specific for fusion proteins comprising stress
PT proteins e.g. heat shock proteins and heterologous proteins e.g. viral
PT antigen, by inducing immune response in host using fusion protein and
PT collecting antibody.
XX
PS Example 1; Fig 4A-B; 29pp; English.
XX
CC This invention describes a novel method for obtaining purified antibodies
CC that specifically bind a fusion protein which comprises a stress protein
CC or its portion, and a heterologous protein or peptide, The method
CC involves introducing a fusion protein into a mammalian host to induce an
CC immune response, removing a sample comprising antibodies from the host
CC and purifying antibodies that specifically bind the fusion protein from
CC the sample. The method is useful for obtaining purified antibodies that
CC specifically bind a fusion protein comprising a stress protein or its
CC portion and a heterologous protein or peptide. The stress protein is a
CC heat shock protein, preferably hsp90, hsp70 or hsp60. The stress protein
CC is a member of the small molecular weight hsp family, preferably a member
CC of groES or Dnaj family, or a human, murine, rat, fungal, parasite or
CC bacterial stress protein. The bacterial stress protein is a member of
CC DnaJ, DnaK, GroES or GroEL stress protein family. The stress protein is a
CC mycobacterial stress protein, preferably a Mycobacterium tuberculosis
CC (71, 65 or 12 kDa protein), M. leprae (70, 65 or 18 kDa protein) or M.
CC bovis Bacille Calmette Guerin (BCG) stress protein (hsp60). The
CC heterologous protein or peptide is a cancer cell-associated antigen, such
CC as GD3, GM2, Gb3, Forssman antigen, Sialosyl-Lea, melanoma antigen,
CC carcino-embryonic antigen (CEA), alpha-fetoprotein, prostate specific
CC antigen or a Tn antigen, a viral antigen such as human immunodeficiency
CC virus (HIV) antigen e.g. HIV gag or pol protein or peptide, an influenza
CC virus antigen such as haemagglutinin, or a parasitic antigen. This
CC sequence represents the Mycobacterium tuberculosis 65kDa stress protein
CC described in the disclosure of the invention
XX
SQ Sequence 540 AA;

Query Match 100.0%; Score 42; DB 6; Length 540;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLLQAAPTL 9
|||||||||
Db 416 TLLQAAPTL 424

RESULT 16
ABU34402
ID ABU34402 standard; protein; 540 AA.
XX
AC ABU34402;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #19929.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Mycobacterium bovis.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR N-PSDB; ACA38272.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 62326; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 540 AA;

Query Match 100.0%; Score 42; DB 6; Length 540;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLLQAAPTL 9
|||||||||
Db 416 TLLQAAPTL 424

RESULT 17
ABU36456
ID ABU36456 standard; protein; 540 AA.
XX
AC ABU36456;
XX
DT 19-JUN-2003 (first entry)

XX
DE Protein encoded by Prokaryotic essential gene #21983.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Mycobacterium tuberculosis.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR N-PSDB; ACA40326.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 64380; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 540 AA;

Query Match 100.0%; Score 42; DB 6; Length 540;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLLQAAPTL 9
|||||||||

Db 416 TLLQAAPTL 424

RESULT 18
ADA24378
ID ADA24378 standard; protein; 540 AA.
XX
AC ADA24378;
XX
DT 20-NOV-2003 (first entry)
XX
DE Mycobacterium tuberculosis 65kDa antigen.
XX
KW Stress protein; immune response; vaccine; mycobacterial stress protein;
KW bacterial stress protein; fungal stress protein; viral stress protein;
KW parasitic stress protein; pathogen; heat shock protein; hsp70; hsp60;
KW p24 protein; immune tolerance; rheumatoid arthritis; autoimmune disease;
KW mycobacterial infection; bacterial infection; viral infection;
KW fungal infection; parasitic infection; immunostimulant; antirheumatic;
KW antiarthritic; antibacterial; virucide; fungicide; antiparasitic;
KW 65kDa antigen.
XX
OS Mycobacterium tuberculosis.
XX
PN US2003073094-A1.
XX
PD 17-APR-2003.
XX
PF 14-JAN-2002; 2002US-00046649.
XX
PR 15-JUN-1988; 88US-00207298.
PR 15-JUN-1989; 89US-00366581.
PR 15-JUN-1989; 89WO-US002619.
PR 09-DEC-1991; 91US-00804632.
PR 04-JUN-1993; 93US-00073381.
PR 06-JUN-1994; 94WO-US006362.
PR 03-NOV-1994; 94US-00336251.
XX
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
XX
PI Young RA, Young D;
XX
DR WPI; 2003-625518/59.
XX
PT New vaccines comprising mycobacterial, bacterial, fungal, viral or
PT parasitic stress proteins, useful for treating, or enhancing an immune
PT response against, e.g. rheumatoid arthritis, or viral or fungal
PT infections, in an individual.
XX
PS Disclosure; Fig 4; 28pp; English.
XX
CC The present invention relates to stress proteins and methods of
CC modulating an individual's immune response. In particular the invention
CC provides a vaccine comprising all, or a portion, of a stress protein, or
CC all or a portion of a protein having an amino acid sequence homologous to
CC the amino acid sequence of the stress protein capable of inducing an
CC immune response in an individual to whom it is administered. Also
CC disclosed is a conjugate comprising a stress protein joined with a
CC substance against which an immune response is desired, a fusion protein
CC comprising a stress protein fused to a protein against which an immune
CC response is desired, and generating antibodies to a substance comprising
CC introducing the conjugate comprised of a stress protein joined to the
CC substance into a mammalian host, removing the antibodies produced by the
CC host to the substance from the host, and purifying the antibodies to
CC generate antibodies to the substance. The stress protein may be selected
CC from mycobacterial stress proteins, bacterial stress proteins, fungal
CC stress proteins, viral stress proteins, or parasitic stress proteins. The
CC stress protein is in sufficient quantity to elicit the desired immune
CC response, or to produce an immune response to the stress protein. In
CC particular, the stress protein is a stress protein of the pathogen. The
CC substance against which an immune response is desired consists of
CC proteins, peptides, oligosaccharides, lipids, carbohydrates, organic
CC molecules, or a combination of these. The vaccine also comprises a

CC recombinant fusion protein, which includes all or a portion of a stress
CC protein, or all or a portion of a protein having an amino acid sequence
CC homologous to the amino acid sequence of the stress protein, fused to a
CC substance against which an immune response is desired or to a portion of
CC the substance sufficient to induce an immune response in an individual.
CC The protein may be the HIV gag or pol protein. The conjugate may comprise
CC a stress protein, which is the heat shock protein 70 (hsp70) or the hsp60
CC protein. The protein may also consist of ovalbumin, influenza virus
CC haemagglutinin protein, human immunodeficiency virus gag protein, or
CC human immunodeficiency virus pol protein. The fusion protein may comprise
CC the stress protein, which is a heat shock protein (particularly hsp70),
CC where the protein is a human immunodeficiency viral protein (specifically
CC p24 protein). The vaccine is useful for enhancing in an individual the
CC immune response to a pathogen, or for producing or enhancing an immune
CC response in an individual. The vaccine is also useful for immunising an
CC individual against subsequent infection by a pathogen, or for inducing in
CC an individual an immune tolerance against a protein (e.g. a protein
CC associated with rheumatoid arthritis). The vaccine is especially useful
CC for treating an autoimmune disease, e.g. rheumatoid arthritis. The
CC vaccine is also useful for treating or preventing mycobacterial,
CC bacterial, viral, fungal, or parasitic infections or diseases. The
CC present sequence represents Mycobacterium tuberculosis 65kDa antigen.
XX
SQ Sequence 540 AA;

Query Match 100.0%; Score 42; DB 6; Length 540;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLLQAAPTL 9
||||||||||
Db 416 TLLQAAPTL 424

RESULT 19
AAW32099
ID AAW32099 standard; protein; 544 AA.
XX
AC AAW32099;
XX
DT 27-AUG-2003 (revised)
DT 08-APR-1998 (first entry)
XX
DE Mycobacteria sp. heat shock protein 65 (hsp65) #1.
XX
KW Heat shock protein; hsp58; human; autoimmune disease; hsp65;
KW rheumatoid arthritis; antigen; infectious disease; prophylactic;
KW pristane induced arthritis; PIA; vaccine.
XX
OS Mycobacteriaceae.
XX
PN WO9711966-A1.
XX
PD 03-APR-1997.
XX
PF 26-SEP-1996; 96WO-GB002382.
XX
PR 27-SEP-1995; 95GB-00019737.
XX
PA (PEPT-) PEPTIDE THERAPEUTIC LTD.
XX
PI Thompson JS, Elson CJ;
XX
DR WPI; 1997-212851/19.
XX
PT Polypeptide(s) derived from microbial heat shock protein - useful for
PT treatment of autoimmune disease esp. arthritis.
XX
PS Disclosure; Page 69-71; 91pp; English.
XX
CC This is the heat shock protein hsp65 which can be used in a novel method
CC to treat autoimmune disease e.g. rheumatoid arthritis. This sequence is
CC known to be an immunodominant antigen in a number of infectious diseases
CC and is linked to pristane induced arthritis (PIA) in vitro. HSP's from
CC microbial sources may act as self antigens and thus have limited use
CC whereas the human hsp65 homologue, hsp58 or fragments of the hsp58
CC protein may be useful in the development of vaccines for prophylaxis or
CC treatment of an autoimmune disease such as rheumatoid arthritis. Note:
CC this sequence is given in the sequence specification but differs slightly
CC from the hsp65 sequence represented in Figure 4 (see AAW32100). (Updated
CC on 27-AUG-2003 to correct OS field.)
XX
SQ Sequence 544 AA;

Query Match 100.0%; Score 42; DB 2; Length 544;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLLQAAPTL 9
||||||||||
Db 420 TLLQAAPTL 428

RESULT 20
AAP80215
ID AAP80215 standard; protein; 560 AA.
XX
AC AAP80215;
XX
DT 25-MAR-2003 (revised)
DT 19-MAR-1991 (first entry)
XX
DE Sequence of Mycobacterium tuberculosis 65kD protein.
XX
KW Antigen; vaccine; ds.
XX
OS Mycobacterium tuberculosis.
XX
PN WO8805823-A.
XX
PD 11-AUG-1988.
XX
PF 01-FEB-1988; 88WO-US000281.
XX
PR 02-FEB-1987; 87US-00010007.
XX
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
XX
PI Husson RN, Young RA, Nick TM;
XX
DR WPI; 1988-235175/33.
DR P-PSDB; AAN80222.
XX
PT Genes encoding Mycobacterium tuberculosis protein antigens - useful for
PT developing reagents for diagnosis, prevention and treatment of
PT tuberculosis.
XX
PS Claim 12; Fig 6; 82pp; English.
XX
CC The gene was isolated by probing a lambda gt11 expression library of
CC M.tuberculosis DNA with monoclonal antibodies directed against
CC M.tuberculosis-specific antigens. 19kD, the 71kD and the 65kD proteins
CC and genes are claimed, and so is a vaccine comprising DNA encoding
CC M.tuberculosis protein in a recombinant vaccine vector. (Updated on 25-
CC MAR-2003 to correct PA field.)
XX
SQ Sequence 560 AA;

Query Match 100.0%; Score 42; DB 1; Length 560;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLLQAAPTL 9
||||||||||
Db 436 TLLQAAPTL 444

```
RESULT 21
AAR04716
ID   AAR04716 standard; protein; 572 AA.
XX
AC   AAR04716;
XX
DT   25-MAR-2003  (revised)
DT   03-JUL-1990  (first entry)
XX
DE   Amino acid sequence of the 65 kDa Mycobacterium tuberculosis protein.
XX
KW   Stress protein; heat shock protein; Mycobacterium tuberculosis protein;
KW   vaccine; human P1 protein; Mycobacterium; immunotherapeutic agent.
XX
OS   Homo sapiens.
XX
PN   WO8912455-A.
XX
PD   28-DEC-1989.
XX
PF   15-JUN-1989;   89WO-US002619.
XX
PR   15-JUN-1988;   88US-00207298.
XX
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PA   (MEDI-) MEDICAL RES COUNCIL.
XX
PI   Young RA,  Young D;
XX
DR   WPI; 1990-022380/03.
XX
PT   Use of stress proteins - for induction or enhancement of immune response
PT   or as and immuno:therapeutic agent in treatment of auto:immune diseases.
XX
PS   Example; Fig 4; 42pp; English.
XX
CC   The stress proteins of Mycobacterium tuberculosis and M. leprae are
CC   involved in humoral and cell-mediated immune responses, and in order to
CC   establish the functions of these proteins, several of the M. tuberculosis
CC   and M. leprae antigens have been sequenced. They exhibit sequence
CC   similarity to known stress-induced proteins. The mycobacterial stress
CC   proteins are recognised by human Abs and T. lymphocytes, and the evidence
CC   suggests that these proteins are among the major targets of the human
CC   cell mediated immune response. Sequence analysis reveals that
CC   mycobacterial hsp60 (alias 65 kDa antigen, groEL) is the major target of
CC   the murine antibody response to both M. tuberculosis and M. leprae.
CC   Comparison of the amino acid sequences of human P1 protein and the 65 kDa
CC   M. tuberculosis protein revealed 257/540 identities. (Updated on 25-MAR-
CC   2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ   Sequence 572 AA;

  Query Match             100.0%;  Score 42;  DB 2;  Length 572;
  Best Local Similarity   100.0%;  Pred. No. 5.1;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TLLQAAPTL 9
              |||||||||
Db        442 TLLQAAPTL 450


RESULT 22
AAR64766
ID   AAR64766 standard; protein; 573 AA.
XX
AC   AAR64766;
XX
DT   25-MAR-2003  (revised)
DT   15-JUL-1995  (first entry)
XX
DE   M. tuberculosis 65 kDa protein.
XX
KW   Immune response; stress protein; mycobacteria; infection response;
KW   T-lymphocyte; vaccine.
XX
OS   Mycobacterium tuberculosis.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 2. .26
FT                   /note= "amino acids at positions 2-26 are not specified"
FT   Misc-difference 177
FT                   /note= "amino acid at position 177 is not specified"
FT   Misc-difference 329
FT                   /note= "amino acid at position 329 is not specified"
FT   Misc-difference 456
FT                   /note= "amino acid at position 456 is not specified"
FT   Misc-difference 560. .564
FT                   /note= "amino acids at positions 560-564 are not
FT                   specified"
XX
PN   WO9429459-A1.
XX
PD   22-DEC-1994.
XX
PF   06-JUN-1994;   94WO-US006362.
XX
PR   04-JUN-1993;   93US-00073381.
XX
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX
PI   Young RA;
XX
DR   WPI; 1995-036486/05.
XX
PT   Use of stress proteins and analogues - for producing or enhancing an
PT   immune response or for inducing immune tolerance, for prophylaxis or
PT   therapy.
XX
PS   Disclosure; Page 43-46; 68pp; English.
XX
CC   The 65 kDa antigens of M. tuberculosis (given in AAR64766) and M. leprae
CC   (EC-AAR64765) are involved in the human T-cell response to mycobacterial
CC   infection. The proteins show homology to the E. coli groEL major stress
CC   protein (AAR64764) and the human P1 protein (AAR64763) and have been used
CC   to construct stress protein-antigen fusion proteins useful as vaccines.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence 573 AA;

  Query Match             100.0%;  Score 42;  DB 2;  Length 573;
  Best Local Similarity   100.0%;  Pred. No. 5.2;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TLLQAAPTL 9
              |||||||||
Db        443 TLLQAAPTL 451


RESULT 23
ABR40246
ID   ABR40246 standard; protein; 583 AA.
XX
AC   ABR40246;
XX
DT   27-JUN-2003  (first entry)
XX
DE   Human recombinant protein vaccine.
XX
KW   Human; vaccine; BCG vaccine heat shock protein 65; prostatic;
KW   antigen cytotoxin T lymphocyte poly epitope; prostate cancer.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   CN1362263-A.
```

XX
PD 07-AUG-2002.
XX
PF 15-NOV-2001; 2001CN-00134935.
XX
PR 04-JAN-2001; 2001CN-00100380.
XX
PA (BEIJ-) BEIJING DEWEIHUAYU BIOTECHNOLOGY CO LTD.
XX
PI Wang L, Li D, Yu Y;
XX
DR WPI; 2003-230415/23.
DR N-PSDB; ACC49833.
XX
PT Recombinant protein vaccine for preventing and treating human prostate
PT cancer.
XX
PS Claim 4; Page 19-22; 35pp; Chinese.
XX
CC The invention relates to a novel recombinant protein vaccine, a fusion
CC protein formed from BCG vaccine heat shock protein 65 and 1-5 copies of
CC human prostatic specific antigen cytotoxin T lymphocyte poly epitope. The
CC vaccine of the invention is useful for treating and preventing carcinoma
CC of prostate. The invention also discloses the genes encoding the
CC recombinant protein vaccines. The present sequence is used in the
CC exemplification of the invention
XX
SQ Sequence 583 AA;

Query Match 100.0%; Score 42; DB 6; Length 583;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLLQAAPTL 9
|||||||||
Db 416 TLLQAAPTL 424

RESULT 24
ABR40247
ID ABR40247 standard; protein; 637 AA.
XX
AC ABR40247;
XX
DT 27-JUN-2003 (first entry)
XX
DE Human recombinant protein vaccine.
XX
KW Human; vaccine; BCG vaccine heat shock protein 65; prostatic;
KW antigen cytotoxin T lymphocyte poly epitope; prostate cancer.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN CN1362263-A.
XX
PD 07-AUG-2002.
XX
PF 15-NOV-2001; 2001CN-00134935.
XX
PR 04-JAN-2001; 2001CN-00100380.
XX
PA (BEIJ-) BEIJING DEWEIHUAYU BIOTECHNOLOGY CO LTD.
XX
PI Wang L, Li D, Yu Y;
XX
DR WPI; 2003-230415/23.
DR N-PSDB; ACC49834.
XX
PT Recombinant protein vaccine for preventing and treating human prostate
PT cancer.
XX
PS Claim 5; Page 27-30; 35pp; Chinese.
XX
CC The invention relates to a novel recombinant protein vaccine, a fusion
CC protein formed from BCG vaccine heat shock protein 65 and 1-5 copies of
CC human prostatic specific antigen cytotoxin T lymphocyte poly epitope. The
CC vaccine of the invention is useful for treating and preventing carcinoma
CC of prostate. The invention also discloses the genes encoding the
CC recombinant protein vaccines. The present sequence is used in the
CC exemplification of the invention
XX
SQ Sequence 637 AA;

Query Match 100.0%; Score 42; DB 6; Length 637;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLLQAAPTL 9
|||||||||
Db 416 TLLQAAPTL 424

RESULT 25
AAB03790
ID AAB03790 standard; protein; 638 AA.
XX
AC AAB03790;
XX
DT 12-SEP-2003 (revised)
DT 06-AUG-2003 (revised)
DT 13-OCT-2000 (first entry)
XX
DE Heat shock protein and tumour specific antigen fusion protein sequence.
XX
KW Heat shock protein; tumour specific antigen; colibacillus; microzyme;
KW plant; immune response; tumour; cancer; human papillomavirus;
KW pointed condyloma.
XX
OS Mycobacterium bovis.
OS Human papillomavirus.
OS Chimeric.
XX
PN CN1248631-A.
XX
PD 29-MAR-2000.
XX
PF 24-SEP-1998; 98CN-00112264.
XX
PR 24-SEP-1998; 98CN-00112264.
XX
PA (ZHOU/) ZHOU G.
XX
PI Zhou G;
XX
DR WPI; 2000-431995/38.
XX
PT New fusion protein for immunotherapy of venereal disease and cancer - is
PT a heat shock protein of Mycobacterium bovis.
XX
PS Claim 9; Fig 1; 5pp; Chinese.
XX
CC The present sequence represents a fusion protein, consisting of a heat
CC shock protein of mycobacterium bovis (var. BCG) fused to a tumour
CC specific antigen of human papillomavirus (HPV). The fusion protein can be
CC expressed in colibacillus, microzymes and plants. The fusion protein is
CC used to make immunostimulant injections, as it can produce specific cell
CC immune and humoral immune responses. It possesses prophylaxis and
CC therapeutic capacity for preventing human papillomavirus (HPV) infection,
CC but also can be used for immunotherapy of pointed condyloma, tumours and
CC cancer caused by HPV. (Updated on 06-AUG-2003 to correct OS field.)
CC (Updated on 12-SEP-2003 to standardise OS field)
XX
SQ Sequence 638 AA;

Query Match 100.0%; Score 42; DB 3; Length 638;

```
  Best Local Similarity   100.0%;  Pred. No. 5.8;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TLLQAAPTL 9
              |||||||||
Db        415 TLLQAAPTL 423


RESULT 26
AAB31609
ID   AAB31609 standard; protein; 639 AA.
XX
AC   AAB31609;
XX
DT   30-APR-2001  (first entry)
XX
DE   Amino acid sequence of Hsp65-E7 fusion protein.
XX
KW   Heat shock protein; Hsp; Th1 response; Th1 cell; CD4+ T lymphocyte cell;
KW   lymphocyte; Hsp65; Hsp40; Hsp10; Hsp60; Hsp71; microbial pathogen;
KW   E7 protein.
XX
OS   Synthetic.
OS   Mycobacterium bovis.
OS   Human papillomavirus.
XX
PN   WO200104344-A2.
XX
PD   18-JAN-2001.
XX
PF   10-JUL-2000; 2000WO-US018828.
XX
PR   08-JUL-1999;   99US-0143757P.
XX
PA   (STRE-) STRESSGEN BIOTECHNOLOGIES CORP.
XX
PI   Siegel M,  Chu NR,  Mizzen LA;
XX
DR   WPI; 2001-138361/14.
DR   N-PSDB; AAF25012.
XX
PT   Screening for compounds that stimulate Th1-like responses in CD4+ T
PT   lymphocyte cells.
XX
PS   Example 5; Fig 4A-B; 88pp; English.
XX
CC   The present sequence represents a fusion protein comprising a
CC   Mycobacterium bovis heat shock protein (Hsp) 65 fused at its C terminal
CC   to a HPV16 E7 protein. The fusion protein is used in the method of the
CC   invention. The specification describes a method of determining whether a
CC   compound stimulates a Th1-like response. Th1 cells are a subset of CD4+ T
CC   lymphocyte cells. The method comprises contacting naive lymphocytes in
CC   vitro with a fusion protein comprising at least a fragment of Hsp, and
CC   then detecting the TH1-like response exhibited by the cell sample. The
CC   proteins which may be used in the method of the invention are Hsp65,
CC   Hsp40, Hsp10, Hsp60, and Hsp71. The method may be used to identify
CC   compounds that stimulate Th1-like responses in response to microbial
CC   pathogens
XX
SQ   Sequence 639 AA;

  Query Match             100.0%;  Score 42;  DB 4;  Length 639;
  Best Local Similarity   100.0%;  Pred. No. 5.8;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TLLQAAPTL 9
              |||||||||
Db        416 TLLQAAPTL 424


RESULT 27
AAB31614
ID   AAB31614 standard; protein; 648 AA.
XX
AC   AAB31614;
XX
DT   30-APR-2001  (first entry)
XX
DE   Amino acid sequence of Hsp65-E7 fusion protein.
XX
KW   Heat shock protein; Hsp; Th1 response; Th1 cell; CD4+ T lymphocyte cell;
KW   lymphocyte; Hsp65; Hsp40; Hsp10; Hsp60; Hsp71; microbial pathogen;
KW   E7 protein.
XX
OS   Synthetic.
OS   Mycobacterium bovis.
OS   Human papillomavirus.
XX
PN   WO200104344-A2.
XX
PD   18-JAN-2001.
XX
PF   10-JUL-2000; 2000WO-US018828.
XX
PR   08-JUL-1999;   99US-0143757P.
XX
PA   (STRE-) STRESSGEN BIOTECHNOLOGIES CORP.
XX
PI   Siegel M,  Chu NR,  Mizzen LA;
XX
DR   WPI; 2001-138361/14.
DR   N-PSDB; AAF25019.
XX
PT   Screening for compounds that stimulate Th1-like responses in CD4+ T
PT   lymphocyte cells.
XX
PS   Example 11; Fig 10A-B; 88pp; English.
XX
CC   The present sequence represents a fusion protein comprising Mycobacterium
CC   bovis heat shock protein (Hsp) 65 fused at its C terminal to a HPV16 E7
CC   protein. The fusion protein is used in the method of the invention. The
CC   specification describes a method of determining whether a compound
CC   stimulates a Th1-like response. Th1 cells are a subset of CD4+ T
CC   lymphocyte cells. The method comprises contacting naive lymphocytes in
CC   vitro with a fusion protein comprising at least a fragment of Hsp, and
CC   then detecting the TH1-like response exhibited by the cell sample. The
CC   proteins which may be used in the method of the invention are Hsp65,
CC   Hsp40, Hsp10, Hsp60, and Hsp71. The method may be used to identify
CC   compounds that stimulate Th1-like responses in response to microbial
CC   pathogens
XX
SQ   Sequence 648 AA;

  Query Match             100.0%;  Score 42;  DB 4;  Length 648;
  Best Local Similarity   100.0%;  Pred. No. 5.9;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TLLQAAPTL 9
              |||||||||
Db        524 TLLQAAPTL 532


RESULT 28
ABG70777
ID   ABG70777 standard; protein; 690 AA.
XX
AC   ABG70777;
XX
DT   06-AUG-2003  (revised)
DT   09-DEC-2002  (first entry)
XX
DE   BCG Hsp65/mutant HBV core antigen, HepCorT(151/97F)Hsp65, protein.
XX
KW   HBV; core antigen; HBe; stress protein; heat shock protein; Hsp65;
KW   immune response; non-cytopathic DNA virus; vaccine; morbidity; mortality;
KW   infection; gene therapy; hepatitis; cirrhosis; hepatocellular carcinoma;
```

```
KW   virucide; immunostimulant; hepatotropic; antiinflammatory; BCG; chimera;
KW   mutant; mutein.
XX
OS   Hepatitis B virus.
OS   Mycobacterium bovis.
OS   Synthetic.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   Region          87. .95
FT                   /note= "Mouse cytotoxic T lymphocyte epitope"
FT   Region          93. .100
FT                   /note= "Murine cytotoxic T lymphocyte epitope"
FT   Misc-difference 97
FT                   /note= "Wild-type Ile substituted by Phe"
XX
PN   WO200262959-A2.
XX
PD   15-AUG-2002.
XX
PF   05-FEB-2002; 2002WO-US003460.
XX
PR   05-FEB-2001; 2001US-0266733P.
XX
PA   (STRE-) STRESSGEN BIOTECHNOLOGIES CORP.
XX
PI   Mizzen L,  Liu H,  Siegel M;
XX
DR   WPI; 2002-706903/76.
DR   N-PSDB; ABS54448.
XX
PT   Novel isolated fusion protein useful for treating hepatitis B virus
PT   infection in a subject, comprises a stress protein or its portion, and a
PT   hepatitis B virus core antigen.
XX
PS   Example 1; Fig 10; 58pp; English.
XX
CC   The invention discloses an isolated fusion protein, and the
CC   polynucleotide encoding it, that comprises a stress protein (e.g.
CC   Mycobacterium bovis heat shock protein, Hsp65), or a portion of it, and a
CC   hepatitis B virus (HBV) core antigen (HBe), where the fusion protein,
CC   when administered to an individual, induces or enhances an immune
CC   response against the HBV core antigen. HBV is a non-cytopathic DNA virus
CC   against which a vaccine has been developed. However, due to the morbidity
CC   and mortality arising from chronic HBV infection occurring over a period
CC   of decades, the impact of vaccinations will not be apparent for some time
CC   yet. The polynucleotide and polypeptide can be used as a vaccine, in gene
CC   therapy and as a pharmaceutical composition for inducing or enhancing an
CC   immune response against an HBV core antigen in a subject for treating an
CC   HBV infection such as hepatitis, cirrhosis and hepatocellular carcinoma.
CC   The sequence presented is the truncated hepatitis B virus core antigen
CC   fused to the N-terminus of the M. bovis variant Bacille-Calmette-Guerin
CC   (BCG) heat shock protein (Hsp) 65, HepCorT(151/97F)Hsp65. The core
CC   antigen has been truncated to include residues 1 to 151 and has had an
CC   amino acid changed from the wild-type sequence. (Updated on 06-AUG-2003
CC   to correct OS field.)
XX
SQ   Sequence 690 AA;

  Query Match             100.0%;  Score 42;  DB 5;  Length 690;
  Best Local Similarity   100.0%;  Pred. No. 6.3;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TLLQAAPTL 9
              |||||||||
Db        566 TLLQAAPTL 574


RESULT 29
ABG70776
ID   ABG70776 standard; protein; 709 AA.
XX
AC   ABG70776;
XX
DT   06-AUG-2003  (revised)
DT   09-DEC-2002  (first entry)
XX
DE   BCG Hsp65/mutant HBV core antigen, hisHepCorT(149/87S97F)Hsp65, protein.
XX
KW   HBV; core antigen; HBe; stress protein; heat shock protein; Hsp65;
KW   immune response; non-cytopathic DNA virus; vaccine; morbidity; mortality;
KW   infection; gene therapy; hepatitis; cirrhosis; hepatocellular carcinoma;
KW   virucide; immunostimulant; hepatotropic; antiinflammatory; BCG; chimera;
KW   mutant; mutein.
XX
OS   Hepatitis B virus.
OS   Mycobacterium bovis.
OS   Synthetic.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .20
FT                   /note= "N-terminus containing a histidine tag"
FT   Region          107. .115
FT                   /note= "Mouse cytotoxic T lymphocyte epitope"
FT   Misc-difference 107
FT                   /note= "Wild-type Asp substituted by Ser"
FT   Region          113. .120
FT                   /note= "Murine cytotoxic T lymphocyte epitope"
FT   Misc-difference 117
FT                   /note= "Wild-type Ile substituted by Phe"
XX
PN   WO200262959-A2.
XX
PD   15-AUG-2002.
XX
PF   05-FEB-2002; 2002WO-US003460.
XX
PR   05-FEB-2001; 2001US-0266733P.
XX
PA   (STRE-) STRESSGEN BIOTECHNOLOGIES CORP.
XX
PI   Mizzen L,  Liu H,  Siegel M;
XX
DR   WPI; 2002-706903/76.
DR   N-PSDB; ABS54447.
XX
PT   Novel isolated fusion protein useful for treating hepatitis B virus
PT   infection in a subject, comprises a stress protein or its portion, and a
PT   hepatitis B virus core antigen.
XX
PS   Example 1; Fig 8; 58pp; English.
XX
CC   The invention discloses an isolated fusion protein, and the
CC   polynucleotide encoding it, that comprises a stress protein (e.g.
CC   Mycobacterium bovis heat shock protein, Hsp65), or a portion of it, and a
CC   hepatitis B virus (HBV) core antigen (HBe), where the fusion protein,
CC   when administered to an individual, induces or enhances an immune
CC   response against the HBV core antigen. HBV is a non-cytopathic DNA virus
CC   against which a vaccine has been developed. However, due to the morbidity
CC   and mortality arising from chronic HBV infection occurring over a period
CC   of decades, the impact of vaccinations will not be apparent for some time
CC   yet. The polynucleotide and polypeptide can be used as a vaccine, in gene
CC   therapy and as a pharmaceutical composition for inducing or enhancing an
CC   immune response against an HBV core antigen in a subject for treating an
CC   HBV infection such as hepatitis, cirrhosis and hepatocellular carcinoma.
CC   The sequence presented is the truncated hepatitis B virus core antigen
CC   fused to the N-terminus of the M. bovis variant Bacille-Calmette-Guerin
CC   (BCG) heat shock protein (Hsp) 65, hisHepCorT(149/87S97F)Hsp65. The core
CC   antigen has been truncated to include residues 1 to 149, had an N-
CC   terminal histidine tag added, had two amino acids changed from the wild-
CC   type sequence and there has been an additional residue added between the
CC   HBV core protein and the Hsp65 protein. (Updated on 06-AUG-2003 to
CC   correct OS field.)
XX
SQ   Sequence 709 AA;
```

```
  Query Match             100.0%;  Score 42;  DB 5;  Length 709;
  Best Local Similarity   100.0%;  Pred. No. 6.5;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TLLQAAPTL 9
              |||||||||
Db        585 TLLQAAPTL 593


RESULT 30
ABG70778
ID   ABG70778 standard; protein; 724 AA.
XX
AC   ABG70778;
XX
DT   06-AUG-2003  (revised)
DT   09-DEC-2002  (first entry)
XX
DE   BCG Hsp65/mutant HBV core antigen, HepCor(97F)Hsp65, protein.
XX
KW   HBV; core antigen; HBe; stress protein; heat shock protein; Hsp65;
KW   immune response; non-cytopathic DNA virus; vaccine; morbidity; mortality;
KW   infection; gene therapy; hepatitis; cirrhosis; hepatocellular carcinoma;
KW   virucide; immunostimulant; hepatotropic; antiinflammatory; BCG; chimera;
KW   mutant; mutein.
XX
OS   Hepatitis B virus.
OS   Mycobacterium bovis.
OS   Synthetic.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   Region          87. .95
FT                   /note= "Mouse cytotoxic T lymphocyte epitope"
FT   Region          93. .100
FT                   /note= "Murine cytotoxic T lymphocyte epitope"
FT   Misc-difference 97
FT                   /note= "Wild-type Ile substituted by Phe"
XX
PN   WO200262959-A2.
XX
PD   15-AUG-2002.
XX
PF   05-FEB-2002; 2002WO-US003460.
XX
PR   05-FEB-2001; 2001US-0266733P.
XX
PA   (STRE-) STRESSGEN BIOTECHNOLOGIES CORP.
XX
PI   Mizzen L,  Liu H,  Siegel M;
XX
DR   WPI; 2002-706903/76.
DR   N-PSDB; ABS54449.
XX
PT   Novel isolated fusion protein useful for treating hepatitis B virus
PT   infection in a subject, comprises a stress protein or its portion, and a
PT   hepatitis B virus core antigen.
XX
PS   Example 1; Fig 12; 58pp; English.
XX
CC   The invention discloses an isolated fusion protein, and the
CC   polynucleotide encoding it, that comprises a stress protein (e.g.
CC   Mycobacterium bovis heat shock protein, Hsp65), or a portion of it, and a
CC   hepatitis B virus (HBV) core antigen (HBe), where the fusion protein,
CC   when administered to an individual, induces or enhances an immune
CC   response against the HBV core antigen. HBV is a non-cytopathic DNA virus
CC   against which a vaccine has been developed. However, due to the morbidity
CC   and mortality arising from chronic HBV infection occurring over a period
CC   of decades, the impact of vaccinations will not be apparent for some time
CC   yet. The polynucleotide and polypeptide can be used as a vaccine, in gene
CC   therapy and as a pharmaceutical composition for inducing or enhancing an
CC   immune response against an HBV core antigen in a subject for treating an
CC   HBV infection such as hepatitis, cirrhosis and hepatocellular carcinoma.
CC   The sequence presented is the hepatitis B virus core antigen fused to the
CC   N-terminus of the M. bovis variant Bacille-Calmette-Guerin (BCG) heat
CC   shock protein (Hsp) 65, HepCor(97F)Hsp65. The core antigen has had an
CC   amino acid changed from the wild-type sequence. (Updated on 06-AUG-2003
CC   to correct OS field.)
XX
SQ   Sequence 724 AA;

  Query Match             100.0%;  Score 42;  DB 5;  Length 724;
  Best Local Similarity   100.0%;  Pred. No. 6.7;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TLLQAAPTL 9
              |||||||||
Db        600 TLLQAAPTL 608


RESULT 31
ABG70775
ID   ABG70775 standard; protein; 746 AA.
XX
AC   ABG70775;
XX
DT   06-AUG-2003  (revised)
DT   09-DEC-2002  (first entry)
XX
DE   His tagged BCG Hsp65/mutant HBV core antigen, hisHepCor(97F)Hsp65.
XX
KW   HBV; core antigen; HBe; stress protein; heat shock protein; Hsp65;
KW   immune response; non-cytopathic DNA virus; vaccine; morbidity; mortality;
KW   infection; gene therapy; hepatitis; cirrhosis; hepatocellular carcinoma;
KW   virucide; immunostimulant; hepatotropic; antiinflammatory; BCG; chimera;
KW   mutant; mutein.
XX
OS   Hepatitis B virus.
OS   Mycobacterium bovis.
OS   Synthetic.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .20
FT                   /note= "N-terminus containing a histidine tag"
FT   Region          107. .115
FT                   /note= "Mouse cytotoxic T lymphocyte epitope"
FT   Region          113. .120
FT                   /note= "Murine cytotoxic T lymphocyte epitope"
FT   Misc-difference 117
FT                   /note= "Wild-type Ile substituted by Phe"
XX
PN   WO200262959-A2.
XX
PD   15-AUG-2002.
XX
PF   05-FEB-2002; 2002WO-US003460.
XX
PR   05-FEB-2001; 2001US-0266733P.
XX
PA   (STRE-) STRESSGEN BIOTECHNOLOGIES CORP.
XX
PI   Mizzen L,  Liu H,  Siegel M;
XX
DR   WPI; 2002-706903/76.
DR   N-PSDB; ABS54446.
XX
PT   Novel isolated fusion protein useful for treating hepatitis B virus
PT   infection in a subject, comprises a stress protein or its portion, and a
PT   hepatitis B virus core antigen.
XX
PS   Example 1; Fig 6; 58pp; English.
XX
CC   The invention discloses an isolated fusion protein, and the
CC   polynucleotide encoding it, that comprises a stress protein (e.g.
```

CC Mycobacterium bovis heat shock protein, Hsp65), or a portion of it, and a
CC hepatitis B virus (HBV) core antigen (HBe), where the fusion protein,
CC when administered to an individual, induces or enhances an immune
CC response against the HBV core antigen. HBV is a non-cytopathic DNA virus
CC against which a vaccine has been developed. However, due to the morbidity
CC and mortality arising from chronic HBV infection occurring over a period
CC of decades, the impact of vaccinations will not be apparent for some time
CC yet. The polynucleotide and polypeptide can be used as a vaccine, in gene
CC therapy and as a pharmaceutical composition for inducing or enhancing an
CC immune response against an HBV core antigen in a subject for treating an
CC HBV infection such as hepatitis, cirrhosis and hepatocellular carcinoma.
CC The sequence presented is the full length hepatitis B virus core antigen
CC fused to the N-terminus of the M. bovis variant Bacille-Calmette-Guerin
CC (BCG) heat shock protein (Hsp) 65, hisHepCor(97F)Hsp65, mutant protein.
CC The core antigen has had an N-terminal histidine tag added and had an
CC amino acid changed from the wild-type sequence. Two additional residues
CC have also been added between the HBV core protein and the Hsp65 protein.
CC (Updated on 06-AUG-2003 to correct OS field.)
XX
SQ Sequence 746 AA;

Query Match 100.0%; Score 42; DB 5; Length 746;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLLQAAPTL 9
|||||||||
Db 622 TLLQAAPTL 630

RESULT 32
AAB31611
ID AAB31611 standard; protein; 948 AA.
XX
AC AAB31611;
XX
DT 30-APR-2001 (first entry)
XX
DE Amino acid sequence of Hsp65-ovalbumin fusion protein.
XX
KW Heat shock protein; Hsp; Th1 response; Th1 cell; CD4+ T lymphocyte cell;
KW lymphocyte; Hsp65; Hsp40; Hsp10; Hsp60; Hsp71; microbial pathogen;
KW ovalbumin.
XX
OS Synthetic.
OS Mycobacterium bovis.
OS Gallus sp.
XX
PN WO200104344-A2.
XX
PD 18-JAN-2001.
XX
PF 10-JUL-2000; 2000WO-US018828.
XX
PR 08-JUL-1999; 99US-0143757P.
XX
PA (STRE-) STRESSGEN BIOTECHNOLOGIES CORP.
XX
PI Siegel M, Chu NR, Mizzen LA;
XX
DR WPI; 2001-138361/14.
DR N-PSDB; AAF25014.
XX
PT Screening for compounds that stimulate Th1-like responses in CD4+ T
PT lymphocyte cells.
XX
PS Example 8; Fig 7A-B; 88pp; English.
XX
CC The present sequence represents a fusion protein comprising a
CC Mycobacterium bovis heat shock protein (Hsp) 65 fused at its C terminal
CC to an ovalbumin protein. The fusion protein is used in the method of the
CC invention. The specification describes a method of determining whether a
CC compound stimulates a Th1-like response. Th1 cells are a subset of CD4+ T
CC lymphocyte cells. The method comprises contacting naive lymphocytes in
CC vitro with a fusion protein comprising at least a fragment of Hsp, and
CC then detecting the TH1-like response exhibited by the cell sample. The
CC proteins which may be used in the method of the invention are Hsp65,
CC Hsp40, Hsp10, Hsp60, and Hsp71. The method may be used to identify
CC compounds that stimulate Th1-like responses in response to microbial
CC pathogens
XX
SQ Sequence 948 AA;

Query Match 100.0%; Score 42; DB 4; Length 948;
Best Local Similarity 100.0%; Pred. No. 8.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLLQAAPTL 9
|||||||||
Db 436 TLLQAAPTL 444

Search completed: October 5, 2004, 07:17:29
Job time : 70 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 5, 2004, 06:07:39 ; Search time 5.09434 Seconds
(without alignments)
91.991 Million cell updates/sec

Title: US-10-022-286-5
Perfect score: 42
Sequence: 1 TLLQAAPTL 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	42	100.0	539	1	CH62_MYCTU	P06806 mycobacteri
2	37	88.1	238	1	RNC_MYCLE	O69469 mycobacteri
3	37	88.1	540	1	CH62_MYCLE	P09239 mycobacteri
4	33	78.6	402	1	PNCB_NEIMA	Q9jtm8 neisseria m
5	33	78.6	1005	1	EVC_MOUSE	P57680 mus musculu
6	32	76.2	240	1	RNC_MYCTU	Q10962 mycobacteri
7	32	76.2	318	1	NU1M_CEBAP	O78695 cebus apell
8	32	76.2	541	1	CH60_NOCFA	Q9afa6 nocardia fa
9	31	73.8	426	1	SYH_PROMM	Q7v4p3 prochloroco
10	30	71.4	239	1	CU58_HUMAN	P58505 homo sapien
11	30	71.4	244	1	LEG3_CRILO	P47953 cricetulus
12	30	71.4	245	1	YI89_COREF	Q8fp96 corynebacte
13	30	71.4	257	1	FAEJ_ECOLI	P33784 escherichia
14	30	71.4	355	1	GLK_SYNY3	Q55855 synechocyst
15	30	71.4	371	1	YKQA_CAEEL	P34306 caenorhabdi
16	30	71.4	402	1	PNCB_NEIMB	Q9jym9 neisseria m
17	30	71.4	476	1	CARL_MOUSE	Q9d5j6 mus musculu
18	30	71.4	503	1	YL32_CAEEL	P34423 caenorhabdi
19	30	71.4	533	1	GLPT_HUMAN	P57057 homo sapien
20	30	71.4	2291	1	SPCB_DROME	Q00963 drosophila
21	29	69.0	142	1	IL3_CALJA	Q28334 callithrix
22	29	69.0	142	1	IL3_SAGOE	P51445 saguinus oe
23	29	69.0	253	1	YDIJ_BACHD	Q9z9p4 bacillus ha
24	29	69.0	254	1	R18B_MOUSE	Q99n84 mus musculu
25	29	69.0	261	1	TESA_MYCTU	Q10974 mycobacteri
26	29	69.0	275	1	HXDC_HUMAN	P35452 homo sapien
27	29	69.0	279	1	HXDC_MOUSE	P23812 mus musculu
28	29	69.0	295	1	KPRS_METMA	Q8pux3 methanosarc
29	29	69.0	306	1	OPPB_HAEIN	P45054 haemophilus
30	29	69.0	317	1	IBP2_SHEEP	Q29400 ovis aries
31	29	69.0	328	1	IBP2_HUMAN	P18065 homo sapien
32	29	69.0	340	1	MMOC_METTR	Q53563 methylosinu
33	29	69.0	378	1	VE2_HPV55	Q80937 human papil
34	29	69.0	406	1	TYRQ_HAEIN	P44747 haemophilus
35	29	69.0	430	1	SYH_SYNPX	Q7u9q6 synechococc
36	29	69.0	434	1	INXE_CAEEL	O62136 caenorhabdi
37	29	69.0	501	1	CP8B_HUMAN	Q9unu6 homo sapien
38	29	69.0	529	1	ATPB_HUMAN	P06576 homo sapien
39	29	69.0	629	1	YD4B_SCHPO	Q10304 schizosacch
40	29	69.0	721	1	ZW10_DROME	Q9w4x9 drosophila
41	29	69.0	732	1	YMM1_CAEEL	P34489 caenorhabdi
42	29	69.0	742	1	COG7_DROME	Q9vad6 drosophila
43	29	69.0	902	1	GLR4_MOUSE	Q9z2w8 mus musculu
44	29	69.0	1278	1	DHBF_BACSU	P45745 bacillus su
45	29	69.0	1780	1	YKZ6_CAEEL	P34333 caenorhabdi

ALIGNMENTS

RESULT 1
CH62_MYCTU
ID CH62_MYCTU STANDARD; PRT; 539 AA.
AC P06806; Q48920; Q48931;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 60 kDa chaperonin 2 (Protein Cpn60-2) (groEL protein 2) (65 kDa
DE antigen) (Heat shock protein 65) (Cell wall protein A) (Antigen A).
GN GROL2 OR GROEL2 OR GROEL-2 OR HSP65 OR RV0440 OR MT0456 OR MTV037.04
GN OR MB0448.
OS Mycobacterium tuberculosis, and
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773, 1765;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=Erdmann;
RX MEDLINE=87137260; PubMed=3029018;
RA Shinnick T.M.;
RT "The 65-kilodalton antigen of Mycobacterium tuberculosis.";
RL J. Bacteriol. 169:1080-1088(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL J. Bacteriol. 184:5479-5490(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=M.bovis; STRAIN=BCG;
RX MEDLINE=87193155; PubMed=3553003;
RA Thole J.E.R., Keulen W.J., Kolk A.H.J., Groothuis D.G., Berwald L.G.,
RA Tiesjema R.H., van Embden J.D.A.;

```
RT   "Characterization, sequence determination, and immunogenicity of a
RT   64-kilodalton protein of Mycobacterium bovis BCG expressed in
RT   Escherichia coli K-12.";
RL   Infect. Immun. 55:1466-1475(1987).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   SPECIES=M.bovis; STRAIN=AF2122/97;
RX   MEDLINE=22709107; PubMed=12788972;
RA   Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA   Pryor M., Duthoy S., Grondin S., Lacroix C., Monsempe C., Simon S.,
RA   Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA   Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT   "The complete genome sequence of Mycobacterium bovis.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
RN   [6]
RP   SEQUENCE OF 45-195 FROM N.A.
RC   SPECIES=M.bovis, and M.tuberculosis; STRAIN=356, and 12-14001;
RA   Ros C., Belak K.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   SEQUENCE OF 63-182 FROM N.A.
RC   SPECIES=M.bovis, and M.tuberculosis; STRAIN=TMC 410, and TMC 1024;
RX   MEDLINE=95150784; PubMed=7848059;
RA   Kapur V., Li L.L., Hamrick M.R., Plikaytis B.B., Shinnick T.M.,
RA   Telenti A., Jacobs W.R. Jr., Banerjee A., Cole S., Yuen K.Y.,
RA   Clarridge J.E., Kreiswirth B.N., Musser J.M.;
RT   "Rapid Mycobacterium species assignment and unambiguous
RT   identification of mutations associated with antimicrobial resistance
RT   in Mycobacterium tuberculosis by automated DNA sequencing.";
RL   Arch. Pathol. Lab. Med. 119:131-138(1995).
RN   [8]
RP   SEQUENCE OF 64-177 FROM N.A.
RC   SPECIES=M.tuberculosis;
RX   MEDLINE=95214306; PubMed=7699930;
RA   Hidaka E., Ueno I., Kawakami Y., Furuwatari C., Furihata K.,
RA   Katsuyama T.;
RT   "Detection and identification of mycobacteria by PCR-RFLP method.";
RL   Rinsho Byori 43:155-161(1995).
CC   -!- FUNCTION: Prevents misfolding and promotes the refolding and
CC       proper assembly of unfolded polypeptides generated under stress
CC       conditions (By similarity).
CC   -!- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of
CC       7 subunits (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- MISCELLANEOUS: PURIFIED 65 kDa ANTIGEN CAN ELICIT A STRONG
CC       DELAYED-TYPE HYPERSENSITIVITY REACTION IN EXPERIMENTAL ANIMALS
CC       INFECTED WITH M.TUBERCULOSIS. THIS PROTEIN IS ONE OF THE MAJOR
CC       IMMUNOREACTIVE PROTEINS OF THE MYCOBACTERIA. THIS ANTIGEN CONTAINS
CC       EPITOPES THAT ARE COMMON TO VARIOUS SPECIES OF MYCOBACTERIA.
CC   -!- SIMILARITY: Belongs to the chaperonin (HSP60) family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M15467; AAA88232.1; -.
DR   EMBL; AL021932; CAA17397.1; -.
DR   EMBL; AE006948; AAK44679.1; -.
DR   EMBL; M17705; AAA25358.1; -.
DR   EMBL; BX248335; CAD93311.1; -.
DR   EMBL; U55833; AAC44451.1; -.
DR   EMBL; U55825; AAC44458.1; -.
DR   EMBL; U17925; AAB39044.1; -.
DR   EMBL; U17957; AAB39076.1; -.
DR   EMBL; S76635; AAP31974.1; -.
DR   PIR; A26950; A26950.
DR   PIR; A43509; A43509.
DR   PIR; S05292; S05292.
DR   HSSP; P06139; 1GRL.
DR   TIGR; MT0456; -.
DR   TubercuList; Rv0440; -.
DR   HAMAP; MF_00600; -; 1.
DR   InterPro; IPR001844; Chaprnin_Cpn60.
DR   InterPro; IPR002423; Cpn60/TCP-1.
DR   InterPro; IPR008950; GroEL-ATPase.
DR   Pfam; PF00118; cpn60_TCP1; 1.
DR   PRINTS; PR00298; CHAPERONIN60.
DR   PRINTS; PR00304; TCOMPLEXTCP1.
DR   PROSITE; PS00296; CHAPERONINS_CPN60; 1.
KW   Chaperone; ATP-binding; Multigene family; Antigen; Cell wall;
KW   Complete proteome.
FT   INIT_MET      0      0       BY SIMILARITY.
SQ   SEQUENCE   539 AA;  56595 MW;  FF03460BAA2BC557 CRC64;

  Query Match             100.0%;  Score 42;  DB 1;  Length 539;
  Best Local Similarity   100.0%;  Pred. No. 0.39;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TLLQAAPTL 9
              |||||||||
Db        415 TLLQAAPTL 423


RESULT 2
RNC_MYCLE
ID   RNC_MYCLE      STANDARD;      PRT;   238 AA.
AC   O69469;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Ribonuclease III (EC 3.1.26.3) (RNase III).
GN   RNC OR ML1659 OR MLCB1243.15.
OS   Mycobacterium leprae.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1769;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TN;
RX   MEDLINE=21128732; PubMed=11234002;
RA   Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
RA   Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA   Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA   Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA   Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA   Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA   Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA   Barrell B.G.;
RT   "Massive gene decay in the leprosy bacillus.";
RL   Nature 409:1007-1011(2001).
CC   -!- FUNCTION: Digests double-stranded RNA. Involved in the processing
CC       of ribosomal RNA precursors and of some mRNAS (By similarity).
CC   -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC       phosphomonoester.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: Contains 1 DRBM (double-stranded RNA-binding) domain.
CC   -!- SIMILARITY: Contains 1 RNase III domain.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AL023635; CAA19196.1; -.
DR   EMBL; AL583923; CAC30612.1; -.
DR   PIR; T44706; T44706.
DR   HSSP; Q91836; 1DI2.
DR   Leproma; ML1659; -.
```

```
DR   HAMAP; MF_00104; -; 1.
DR   InterPro; IPR001159; DS_RBD.
DR   InterPro; IPR000999; RNase_III.
DR   Pfam; PF00035; dsrm; 1.
DR   Pfam; PF00636; Ribonuclease_3; 1.
DR   SMART; SM00358; DSRM; 1.
DR   SMART; SM00535; RIBOc; 1.
DR   PROSITE; PS50137; DS_RBD; 1.
DR   PROSITE; PS00517; RNASE_3_1; 1.
DR   PROSITE; PS50142; RNASE_3_2; 1.
KW   Hydrolase; Nuclease; Endonuclease; RNA-binding; Complete proteome.
FT   DOMAIN       19    134       RNASE III.
FT   DOMAIN      211    227       DRBM.
SQ   SEQUENCE   238 AA;  25240 MW;  84BC036362C92AFE CRC64;

  Query Match             88.1%;  Score 37;  DB 1;  Length 238;
  Best Local Similarity   88.9%;  Pred. No. 1.9;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TLLQAAPTL 9
              ||| |||||
Db        148 TLLDAAPTL 156


RESULT 3
CH62_MYCLE
ID   CH62_MYCLE     STANDARD;      PRT;   540 AA.
AC   P09239;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   60 kDa chaperonin 2 (Protein Cpn60 2) (groEL protein 2) (65 kDa
DE   antigen).
GN   GROL2 OR GROEL2 OR GROEL-2 OR ML0317 OR MLCB1450.05C.
OS   Mycobacterium leprae.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1769;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86313701; PubMed=2428046;
RA   Mehra V., Sweetser D., Young R.A.;
RT   "Efficient mapping of protein antigenic determinants.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:7013-7017(1986).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TN;
RX   MEDLINE=21128732; PubMed=11234002;
RA   Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
RA   Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA   Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA   Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA   Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA   Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA   Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA   Barrell B.G.;
RT   "Massive gene decay in the leprosy bacillus.";
RL   Nature 409:1007-1011(2001).
CC   -!- FUNCTION: Prevents misfolding and promotes the refolding and
CC       proper assembly of unfolded polypeptides generated under stress
CC       conditions (By similarity).
CC   -!- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of
CC       7 subunits (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- MISCELLANEOUS: PURIFIED 65 KDA ANTIGEN CAN ELICIT A STRONG
CC       DELAYED-TYPE HYPERSENSITIVITY REACTION IN EXPERIMENTAL ANIMALS
CC       INFECTED WITH MYCOBACTERIUM TUBERCULOSIS.
CC   -!- MISCELLANEOUS: This protein is one of the major immunoreactive
CC       proteins of the Mycobacteria. This antigen contains epitopes that
CC       are common to various species of Mycobacteria.
CC   -!- SIMILARITY: Belongs to the chaperonin (HSP60) family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M14341; AAA25354.1; ALT_INIT.
DR   EMBL; AL035159; CAA22689.1; -.
DR   EMBL; AL583918; CAC29825.1; -.
DR   PIR; A25902; A25902.
DR   PIR; T44725; T44725.
DR   HSSP; P06139; 1GRL.
DR   Leproma; ML0317; -.
DR   HAMAP; MF_00600; -; 1.
DR   InterPro; IPR001844; Chaprnin_Cpn60.
DR   InterPro; IPR002423; Cpn60/TCP-1.
DR   InterPro; IPR008950; GroEL-ATPase.
DR   Pfam; PF00118; cpn60_TCP1; 1.
DR   PRINTS; PR00298; CHAPERONIN60.
DR   PRINTS; PR00304; TCOMPLEXTCP1.
DR   PROSITE; PS00296; CHAPERONINS_CPN60; 1.
KW   Chaperone; Antigen; ATP-binding; Multigene family; Complete proteome.
FT   INIT_MET     0      0       BY SIMILARITY.
SQ   SEQUENCE   540 AA;  56760 MW;  731498AFBA3A2AEF CRC64;

  Query Match             88.1%;  Score 37;  DB 1;  Length 540;
  Best Local Similarity   88.9%;  Pred. No. 4.4;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TLLQAAPTL 9
              |||||||  |
Db        415 TLLQAAPAL 423


RESULT 4
PNCB_NEIMA
ID   PNCB_NEIMA     STANDARD;      PRT;   402 AA.
AC   Q9JTM8;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Nicotinate phosphoribosyltransferase (EC 2.4.2.11) (NAPRTase).
GN   PNCB OR NMA1706.
OS   Neisseria meningitidis (serogroup A).
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=65699;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Z2491 / Serogroup A / Serotype 4A;
RX   MEDLINE=20222556; PubMed=10761919;
RA   Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA   Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA   Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA   Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA   Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA   Whitehead S., Spratt B.G., Barrell B.G.;
RT   "Complete DNA sequence of a serogroup A strain of Neisseria
RT   meningitidis Z2491.";
RL   Nature 404:502-506(2000).
CC   -!- CATALYTIC ACTIVITY: Nicotinate D-ribonucleotide + diphosphate =
CC       nicotinate + 5-phospho-alpha-D-ribose 1-diphosphate.
CC   -!- PATHWAY: NAD biosynthesis; nicotinamide to NaMN; second step.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: Belongs to the NAPRTase family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
```

```
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AL162756; CAB84934.1; ALT_INIT.
DR   HAMAP; MF_00570; -; 1.
DR   InterPro; IPR007229; NAPRTase.
DR   InterPro; IPR006406; Nic_Prtrans.
DR   Pfam; PF04095; NAPRTase; 1.
DR   TIGRFAMs; TIGR01514; NAPRTase; 1.
KW   Pyridine nucleotide biosynthesis; Transferase; Glycosyltransferase;
KW   Complete proteome.
SQ   SEQUENCE   402 AA;  46340 MW;  92A02C6F588858B6 CRC64;

  Query Match             78.6%;  Score 33;  DB 1;  Length 402;
  Best Local Similarity   66.7%;  Pred. No. 22;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TLLQAAPTL 9
              |||:|||::
Db        193 TLLEAAPSI 201


RESULT 5
EVC_MOUSE
ID   EVC_MOUSE      STANDARD;      PRT;  1005 AA.
AC   P57680;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Ellis-van Creveld syndrome protein homolog.
GN   EVC.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC   TISSUE=Brain;
RX   MEDLINE=20164328; PubMed=10700184;
RA   Ruiz-Perez V.L., Ide S.E., Strom T.M., Lorenz B., Wilson D., Woods K.,
RA   King L., Francomano C., Freisinger P., Spranger S., Marino B.,
RA   Dallapiccola B., Wright M., Meitinger T., Polymeropoulos M.H.,
RA   Goodship J.;
RT   "Mutations in a new gene in Ellis-van Creveld syndrome and Weyers
RT   acrodental dysostosis.";
RL   Nat. Genet. 24:283-286(2000).
RN   [2]
RP   ERRATUM.
RA   Ruiz-Perez V.L., Ide S.E., Strom T.M., Lorenz B., Wilson D., Woods K.,
RA   King L., Francomano C., Freisinger P., Spranger S., Marino B.,
RA   Dallapiccola B., Wright M., Meitinger T., Polymeropoulos M.H.,
RA   Goodship J.;
RL   Nat. Genet. 25:125-125(2000).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=Long;
CC         IsoId=P57680-1; Sequence=Displayed;
CC       Name=Short;
CC         IsoId=P57680-2; Sequence=VSP_004247;
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AJ250841; CAB76567.1; -.
DR   EMBL; AJ250841; CAB76568.1; ALT_SEQ.
DR   MGD; MGI:1890596; Evc.
KW   Transmembrane; Alternative splicing.
FT   TRANSMEM     22     42       POTENTIAL.
FT   VARSPLIC    966    978       Missing (in isoform Short).
FT                                /FTId=VSP_004247.
SQ   SEQUENCE   1005 AA;  113003 MW;  D69F25ABB443BD0E CRC64;

  Query Match             78.6%;  Score 33;  DB 1;  Length 1005;
  Best Local Similarity   100.0%;  Pred. No. 56;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LQAAPTL 9
              |||||||
Db         16 LQAAPTL 22


RESULT 6
RNC_MYCTU
ID   RNC_MYCTU      STANDARD;      PRT;   240 AA.
AC   Q10962;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ribonuclease III (EC 3.1.26.3) (RNase III).
GN   RNC OR RV2925C OR MT2995 OR MTCY338.14C OR MB2950C.
OS   Mycobacterium tuberculosis, and
OS   Mycobacterium bovis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1773, 1765;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   SPECIES=M.tuberculosis; STRAIN=H37Rv;
RX   MEDLINE=98295987; PubMed=9634230;
RA   Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA   Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F.,
RA   Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA   Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA   Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA   Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA   Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the
RT   complete genome sequence.";
RL   Nature 393:537-544(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;
RX   MEDLINE=22206494; PubMed=12218036;
RA   Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA   Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA   Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA   Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA   Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
RT   "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT   laboratory strains.";
RL   J. Bacteriol. 184:5479-5490(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   SPECIES=M.bovis; STRAIN=AF2122/97;
RX   MEDLINE=22709107; PubMed=12788972;
RA   Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA   Pryor M., Duthoy S., Grondin S., Lacroix C., Monsempe C., Simon S.,
RA   Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA   Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT   "The complete genome sequence of Mycobacterium bovis.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
CC   -!- FUNCTION: Digests double-stranded RNA. Involved in the processing
CC       of ribosomal RNA precursors and of some mRNAS (By similarity).
CC   -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC       phosphomonoester.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: Contains 1 DRBM (double-stranded RNA-binding) domain.
CC   -!- SIMILARITY: Contains 1 RNase III domain.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; Z74697; CAA98974.1; -.
DR   EMBL; AE007121; AAK47322.1; -.
DR   EMBL; BX248344; CAD96637.1; -.
DR   PIR; E70748; E70748.
DR   HSSP; Q91836; 1DI2.
DR   TIGR; MT2995; -.
DR   TubercuList; Rv2925c; -.
DR   HAMAP; MF_00104; -; 1.
DR   InterPro; IPR001159; DS_RBD.
DR   InterPro; IPR000999; RNase_III.
DR   Pfam; PF00035; dsrm; 1.
DR   Pfam; PF00636; Ribonuclease_3; 1.
DR   SMART; SM00358; DSRM; 1.
DR   SMART; SM00535; RIBOc; 1.
DR   PROSITE; PS50137; DS_RBD; 1.
DR   PROSITE; PS00517; RNASE_3_1; 1.
DR   PROSITE; PS50142; RNASE_3_2; 1.
KW   Hydrolase; Nuclease; Endonuclease; RNA-binding; Complete proteome.
FT   DOMAIN       19    134       RNASE III.
FT   DOMAIN      211    227       DRBM.
SQ   SEQUENCE   240 AA;  25399 MW;  64C8636742F161DC CRC64;

  Query Match             76.2%;  Score 32;  DB 1;  Length 240;
  Best Local Similarity   87.5%;  Pred. No. 22;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 LLQAAPTL 9
              || |||||
Db        149 LLDAAPTL 156


RESULT 7
NU1M_CEBAP
ID   NU1M_CEBAP     STANDARD;      PRT;   318 AA.
AC   O78695;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   NADH-ubiquinone oxidoreductase chain 1 (EC 1.6.5.3).
GN   MTND1 OR ND1.
OS   Cebus apella (Brown-capped capuchin).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae; Cebus.
OX   NCBI_TaxID=9515;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Muscle;
RX   MEDLINE=98404151; PubMed=9732458;
RA   Cao Y., Janke A., Waddell P.J., Westerman M., Takenaka O., Murata S.,
RA   Okada N., Paabo S., Hasegawa M.;
RT   "Conflict among individual mitochondrial proteins in resolving the
RT   phylogeny of eutherian orders.";
RL   J. Mol. Evol. 47:307-322(1998).
CC   -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC   -!- SIMILARITY: Belongs to the complex I subunit 1 family.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AB010973; BAA32098.1; -.
DR   InterPro; IPR001694; Resp_NADH_dh1.
DR   Pfam; PF00146; NADHdh; 1.
DR   PROSITE; PS00667; COMPLEX1_ND1_1; 1.
DR   PROSITE; PS00668; COMPLEX1_ND1_2; 1.
KW   Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transmembrane.
SQ   SEQUENCE   318 AA;  35805 MW;  D8BA793FE59AF00A CRC64;

  Query Match             76.2%;  Score 32;  DB 1;  Length 318;
  Best Local Similarity   77.8%;  Pred. No. 29;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TLLQAAPTL 9
              ||  |||||
Db         69 TLYMAAPTL 77


RESULT 8
CH60_NOCFA
ID   CH60_NOCFA     STANDARD;      PRT;   541 AA.
AC   Q9AFA6;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   60 kDa chaperonin (Protein Cpn60) (groEL protein) (Heat shock protein
DE   60).
GN   GROL OR GROEL OR HSP60.
OS   Nocardia farcinica.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Nocardiaceae; Nocardia.
OX   NCBI_TaxID=37329;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Zimmermann O.S., Koechel H.G.;
RT   "Nocardia farcinica heat shock protein 60 (hsp60) gene.";
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Prevents misfolding and promotes the refolding and
CC       proper assembly of unfolded polypeptides generated under stress
CC       conditions (By similarity).
CC   -!- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of
CC       7 subunits (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: Belongs to the chaperonin (HSP60) family.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AF352577; AAK18613.1; -.
DR   HSSP; P06139; 1GRL.
DR   HAMAP; MF_00600; -; 1.
DR   InterPro; IPR001844; Chaprnin_Cpn60.
DR   InterPro; IPR002423; Cpn60/TCP-1.
DR   InterPro; IPR008950; GroEL-ATPase.
DR   Pfam; PF00118; cpn60_TCP1; 1.
DR   PRINTS; PR00298; CHAPERONIN60.
DR   PRINTS; PR00304; TCOMPLEXTCP1.
DR   PROSITE; PS00296; CHAPERONINS_CPN60; 1.
KW   Chaperone; ATP-binding.
SQ   SEQUENCE   541 AA;  56480 MW;  7980C3E06159CF5B CRC64;

  Query Match             76.2%;  Score 32;  DB 1;  Length 541;
  Best Local Similarity   87.5%;  Pred. No. 49;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 LLQAAPTL 9
              |||||| |
Db        417 LLQAAPAL 424
```

```
RESULT 9
SYH_PROMM
ID   SYH_PROMM      STANDARD;      PRT;   426 AA.
AC   Q7V4P3;
DT   15-MAR-2004 (Rel. 43, Created)
DT   15-MAR-2004 (Rel. 43, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Histidyl-tRNA synthetase (EC 6.1.1.21) (Histidine--tRNA ligase)
DE   (HisRS).
GN   HISS OR PMT1905.
OS   Prochlorococcus marinus (strain MIT 9313).
OC   Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC   Prochlorococcus.
OX   NCBI_TaxID=74547;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22825698; PubMed=12917642;
RA   Rocap G., Larimer F.W., Lamerdin J., Malfatti S., Chain P.,
RA   Ahlgren N.A., Arellano A., Coleman M., Hauser L., Hess W.R.,
RA   Johnson Z.I., Land M., Lindell D., Post A.F., Regala W., Shah M.,
RA   Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,
RA   Webb E.A., Zinser E.R., Chisholm S.W.;
RT   "Genome divergence in two Prochlorococcus ecotypes reflects oceanic
RT   niche differentiation.";
RL   Nature 424:1042-1047(2003).
CC   -!- CATALYTIC ACTIVITY: ATP + L-histidine + tRNA(His) = AMP +
CC       diphosphate + L-histidyl-tRNA(His).
CC   -!- SUBUNIT: Homodimer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; BX572100; CAE22080.1; -.
DR   HAMAP; MF_00127; -; 1.
DR   InterPro; IPR004154; HGTP_anticodon.
DR   InterPro; IPR004516; HisS.
DR   InterPro; IPR002314; tRNA-synt_2b.
DR   InterPro; IPR006195; tRNA_ligase_II.
DR   Pfam; PF03129; HGTP_anticodon; 1.
DR   Pfam; PF00587; tRNA-synt_2b; 1.
DR   TIGRFAMs; TIGR00442; hisS; 1.
DR   PROSITE; PS50862; AA_TRNA_LIGASE_II; 1.
KW   Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW   Complete proteome.
SQ   SEQUENCE   426 AA;  47360 MW;  5DFDB70AC6B6DBC2 CRC64;

  Query Match             73.8%;  Score 31;  DB 1;  Length 426;
  Best Local Similarity   87.5%;  Pred. No. 62;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 LLQAAPTL 9
              ||| ||||
Db        206 LLQDAPTL 213


RESULT 10
CU58_HUMAN
ID   CU58_HUMAN     STANDARD;      PRT;   239 AA.
AC   P58505;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Protein C21orf58.
GN   C21ORF58.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS A AND B).
RX   MEDLINE=21564202; PubMed=11707072;
RA   Reymond A., Friedli M., Neergaard Henrichsen C., Chapot F.,
RA   Deutsch S., Ucla C., Rossier C., Lyle R., Guipponi M.,
RA   Antonarakis S.E.;
RT   "From PREDs and open reading frames to cDNA isolation: revisiting the
RT   human chromosome 21 transcription map.";
RL   Genomics 78:46-54(2001).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=A;
CC         IsoId=P58505-1; Sequence=Displayed;
CC       Name=B;
CC         IsoId=P58505-2; Sequence=VSP_003827, VSP_003828;
CC   -!- TISSUE SPECIFICITY: Expressed in skin and fetal lung.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AY039243; AAK72408.1; -.
DR   EMBL; AY039244; AAK72409.1; -.
DR   Genew; HGNC:1300; C21orf58.
KW   Alternative splicing.
FT   DOMAIN      209    216       POLY-HIS.
FT   VARSPLIC    190    194       PPHVP -> RFFHS (in isoform B).
FT                                /FTId=VSP_003827.
FT   VARSPLIC    195    239       Missing (in isoform B).
FT                                /FTId=VSP_003828.
SQ   SEQUENCE   239 AA;  26192 MW;  F40E9244608D567A CRC64;

  Query Match             71.4%;  Score 30;  DB 1;  Length 239;
  Best Local Similarity   66.7%;  Pred. No. 57;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TLLQAAPTL 9
              |:|| ||:|
Db        225 TVLQPAPSL 233


RESULT 11
LEG3_CRILO
ID   LEG3_CRILO     STANDARD;      PRT;   244 AA.
AC   P47953;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Galectin-3 (Galactose-specific lectin 3) (MAC-2 antigen) (IgE-binding
DE   protein) (35 kDa lectin) (Carbohydrate binding protein 35) (CBP 35)
DE   (Laminin-binding protein) (Lectin L-29) (CBP30).
GN   LGALS3.
OS   Cricetulus longicaudatus (Long-tailed hamster) (Chinese hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Cricetulus.
OX   NCBI_TaxID=10030;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=94299546; PubMed=8027086;
RA   Mehul B., Bawumia S., Martin S.R., Hughes R.C.;
RT   "Structure of baby hamster kidney carbohydrate-binding protein CBP30,
RT   an S-type animal lectin.";
RL   J. Biol. Chem. 269:18250-18258(1994).
CC   -!- FUNCTION: Galactose-specific lectin which binds IgE.
CC   -!- SIMILARITY: In the C-terminal section; belongs to the galectin
```

```
CC       (galaptin/S-lectin) family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X78879; CAA55479.1; -.
DR   HSSP; P17931; 1A3K.
DR   InterPro; IPR008985; ConA_like_lec_gl.
DR   InterPro; IPR001079; Galectin.
DR   Pfam; PF00337; Gal-bind_lectin; 1.
DR   SMART; SM00276; GLECT; 1.
DR   PROSITE; PS00309; GALAPTIN; 1.
KW   Galectin; Lectin; IgE-binding protein; Repeat; Phosphorylation;
KW   Acetylation.
FT   INIT_MET      0      0       BY SIMILARITY.
FT   MOD_RES       1      1       ACETYLATION (BY SIMILARITY).
FT   MOD_RES       5      5       PHOSPHORYLATION (BY CK1)
FT                                (BY SIMILARITY).
FT   DOMAIN       34     98       7 X 9 AA TANDEM REPEATS OF Y-P-G-X(3)-P-
FT                                [GS]-A.
FT   REPEAT       34     42       1.
FT   REPEAT       43     51       2.
FT   REPEAT       52     60       3.
FT   REPEAT       61     69       4.
FT   REPEAT       70     77       5 (APPROXIMATE).
FT   REPEAT       78     87       6 (APPROXIMATE).
FT   REPEAT       88     98       7 (APPROXIMATE).
FT   DOMAIN      112    244       GALAPTIN.
FT   DISULFID    167    167       INTERCHAIN (BY SIMILARITY).
FT   BINDING     175    181       BETA-GALACTOSIDE (BY SIMILARITY).
SQ   SEQUENCE   244 AA;  25608 MW;  8F99B9AA0BBA7D3F CRC64;

  Query Match             71.4%;  Score 30;  DB 1;  Length 244;
  Best Local Similarity   66.7%;  Pred. No. 58;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TLLQAAPTL 9
              ||  ||||:
Db        235 TLTSAAPTM 243


RESULT 12
YI89_COREF
ID   YI89_COREF     STANDARD;      PRT;   245 AA.
AC   Q8FP96;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Hypothetical UPF0246 protein CE1889.
GN   CE1889.
OS   Corynebacterium efficiens.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX   NCBI_TaxID=152794;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
RX   MEDLINE=22723752; PubMed=12840036;
RA   Nishio Y., Nakamura Y., Kawarabayasi Y., Usuda Y., Kimura E.,
RA   Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,
RA   Gojobori T.;
RT   "Comparative complete genome sequence analysis of the amino acid
RT   replacements responsible for the thermostability of Corynebacterium
RT   efficiens.";
RL   Genome Res. 13:1572-1579(2003).
CC   -!- SIMILARITY: Belongs to the UPF0246 family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AP005220; BAC18699.1; ALT_INIT.
DR   HAMAP; MF_00652; -; 1.
DR   InterPro; IPR005583; DUF328.
DR   Pfam; PF03883; DUF328; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   245 AA;  25594 MW;  D11BE88E5A5C54BE CRC64;

  Query Match             71.4%;  Score 30;  DB 1;  Length 245;
  Best Local Similarity   55.6%;  Pred. No. 58;
  Matches    5;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TLLQAAPTL 9
              |:|::|||:
Db         67 TVLRSAPTM 75


RESULT 13
FAEJ_ECOLI
ID   FAEJ_ECOLI     STANDARD;      PRT;   257 AA.
AC   P33784;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   K88 minor fimbrial subunit faeJ precursor.
GN   FAEJ.
OS   Escherichia coli.
OG   Plasmid pFM205.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93015683; PubMed=1400188;
RA   Bakker D., Willemsen P.T.J., Willems R.H., Huisman T.T., Mooi F.R.,
RA   Oudega B., Stegehuis F., de Graaf F.K.;
RT   "Identification of minor fimbrial subunits involved in biosynthesis
RT   of K88 fimbriae.";
RL   J. Bacteriol. 174:6350-6358(1992).
CC   -!- FUNCTION: K88 MINOR FIMBRIAL SUBUNIT, PLAYS AN ESSENTIAL
CC       ROLE IN THE BIOGENESIS OF THE K88 FIMBRIAE. FIMBRIAE (ALSO CALLED
CC       PILI), POLAR FILAMENTS RADIATING FROM THE SURFACE OF THE BACTERIUM
CC       TO A LENGTH OF 0.5-1.5 MICROMETERS AND NUMBERING 100-300 PER CELL,
CC       ENABLE BACTERIA TO COLONIZE THE EPITHELIUM OF SPECIFIC HOST
CC       ORGANS.
CC   -!- SUBCELLULAR LOCATION: Located in or along the K88 fimbrial
CC       structure.
CC   -!- SIMILARITY: TO THE MAJOR K88 FIMBRIAL SUBUNIT (FAEG).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z11700; CAA77762.1; -.
DR   EMBL; Z11710; CAA77773.1; -.
DR   PIR; S24812; S24812.
KW   Fimbria; Signal; Plasmid.
FT   SIGNAL        1     26       POTENTIAL.
FT   CHAIN        27    257       K88 MINOR FIMBRIAL SUBUNIT FAEJ.
SQ   SEQUENCE   257 AA;  27954 MW;  0FFD2AB260BCE7C8 CRC64;

  Query Match             71.4%;  Score 30;  DB 1;  Length 257;
  Best Local Similarity   75.0%;  Pred. No. 61;
```

```
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TLLQAAPT 8
              |||| ||:
Db        116 TLLQVAPS 123


RESULT 14
GLK_SYNY3
ID   GLK_SYNY3      STANDARD;      PRT;   355 AA.
AC   Q55855;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Glucokinase (EC 2.7.1.2) (Glucose kinase).
GN   GLK OR SLL0593.
OS   Synechocystis sp. (strain PCC 6803).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX   NCBI_TaxID=1148;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96127529; PubMed=8590279;
RA   Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA   Sugiura M., Tabata S.;
RT   "Sequence analysis of the genome of the unicellular cyanobacterium
RT   Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT   region from map positions 64% to 92% of the genome.";
RL   DNA Res. 2:153-166(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97061201; PubMed=8905231;
RA   Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA   Miyajima N., Hirosawa M., Sugiura M., Sasamoto S., Kimura T.,
RA   Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA   Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA   Tabata S.;
RT   "Sequence analysis of the genome of the unicellular cyanobacterium
RT   Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT   entire genome and assignment of potential protein-coding regions.";
RL   DNA Res. 3:109-136(1996).
CC   -!- CATALYTIC ACTIVITY: ATP + D-glucose = ADP + D-glucose 6-phosphate.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: Belongs to the bacterial glucokinase family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D64004; BAA10611.1; -.
DR   PIR; S76667; S76667.
DR   HAMAP; MF_00524; -; 1.
DR   InterPro; IPR003836; Glucokinase.
DR   Pfam; PF02685; Glucokinase; 1.
DR   TIGRFAMs; TIGR00749; glk; 1.
KW   Transferase; Kinase; Glycolysis; ATP-binding; Complete proteome.
FT   NP_BIND      11      16       ATP (POTENTIAL).
SQ   SEQUENCE   355 AA;  37712 MW;  01E16E363ED820E6 CRC64;

  Query Match             71.4%;  Score 30;  DB 1;  Length 355;
  Best Local Similarity   85.7%;  Pred. No. 84;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TLLQAAP 7
              |:|||||
Db        133 TVLQAAP 139


RESULT 15
YKQA_CAEEL
ID   YKQA_CAEEL     STANDARD;      PRT;   371 AA.
AC   P34306;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hypothetical protein C06E1.11 in chromosome III.
GN   C06E1.11.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=94150718; PubMed=7906398;
RA   Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA   Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA   Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA   Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA   Johnston L., Jones M., Kershaw J., Kirsten J., Laisster N.,
RA   Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA   Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA   Sims M., Smaldon N., Smith A., Smith M., Sonnhammer E., Staden R.,
RA   Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA   Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA   Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L16559; AAA27927.1; -.
DR   PIR; A88534; A88534.
DR   WormPep; C06E1.11; CE00056.
KW   Hypothetical protein.
SQ   SEQUENCE   371 AA;  38544 MW;  AC64458BD29F6E52 CRC64;

  Query Match             71.4%;  Score 30;  DB 1;  Length 371;
  Best Local Similarity   66.7%;  Pred. No. 88;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TLLQAAPTL 9
              |: |||| |
Db        241 TIAQAAPVL 249


Search completed: October  5, 2004, 06:17:34
Job time : 8.09434 secs
```